

FIGURE 1

GGGCCTCGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTACAAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTGTGAAGTTGTAACAGAAAACCTGTTAGAAATGTGGTGGT
TTCAGCAAGGCCTCAGTTCCCTCCTCAGCCCTTGTAAATTGGACATCTGCTGCTTCATATT
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTACCTTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTGGGCAATGCTAAATATTGCGGCAGTTATGCATTG
CTACCATTATGTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAACGTTATCATCAA
TTAAACAAGGCTGGCCTTGTACTTGAATACTGAGTTAGGACTTCTATTGTGGCAAACCTT
CCAGAAAACAACCCTTTGCTGCACATGTAAGTGGAGCTGTGCTTACCTTGGTATGGCTCAT
TATATATGTTGTTCAGACCATCCTTCTACCAAATGCAGCCAAAATCCATGGCAAACAGTC
TTCTGGATCAGACTGTTGGTTATCTGGTGTGGAGTAAGTGCACTTAGCATGCTGACTTGCTC
ATCAGTTTGACAGTGGCAATTGGACTGATTAGAACAGAAACTCCATTGGAACCCGAGG
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAACATGGCTATGTCATTTCCTTCTT
GGTTTTCTGACTTACATTGATGTTTCAAGAAAATTCTTACGGGTGGAAGCCAATTACA
TGGATTAACCTCTATGACACTGCACCTTGCCTTATTAAACATGAACGAACACGGCTACTTCCA
GAGATATTTGATGAAGGATAAAATTTCTGTAATGATTATGATTCTCAGGGATTGGGAAAGG
TTCACAGAAGTTGCTTATTCTCTGAAATTCAACCCTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAAACATGAAAGAACCCATTGATAGATTCTAAAGGATATCAT
CAAGAAGACTATTAAAAACACCTATGCCTATACTTTTATCTCAGAAAATAAGTCAAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIFSYYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV
LCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAVLTFG
MGSLYMFVQTILSYQMOPKIHGKQFWIRLLLVIWCGVSALSMLTCSSVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAEWSMSFSFFGFFLTYIRDFQKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCGTGGCGGACCGTGGGGAGAGCCGCAGTCGGCTGCAGCACCTGGAGAAGGCAGACC
GTGTGAGGGGGCTGTGGCCCCAGCGTGTGGCTCGGGAGTGGAAAGTGGAGGCAGGAGCCTTC
CTTACACTTCGCC**CATGAGTT**CCTCATCGACTCCAGCATCATGATTACCTCCAGATACTATTTTG
GATTGGGTGGCTTTCTTCATGCCCAATTGTTAAAGACTATGAGATACTGTCAGTATGTTGTACAG
GTGATCTCTCCGTGACGTTGCATTTCTGCACCATGTTGAGCTCATCATCTTGAAATCTTAGG
AGTATTGAATAGCAGCTCCGTTATTCACTGGAAAATGAACCTGTGTAACTCTGCTGATCCTGG
TTTCATGGTGCCTTTACATTGGCTATTTATTGTGAGCAATATCCGACTACTGCATAAACACGA
CTGCTTTTCCTGTCTTATGGCTGACCTTATGTATTTCTTGAAACTAGGAGATCCCTTCC
CATTCTCAGCCCCAAACATGGGATCTTACATAGAACAGCTCATCAGCCGGGTGGTGTGATTGGAG
TGACTCTCATGGCTCTCTGGATTGGTGTCAACTGCCATACACTACATGTCTTACTTC
CTCAGGAATGTGACTGACACGGATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATATGAT
CATAGAAAAAGAAAAGGATGGCAATGGCACGGAGAACATGTTCCAGAAGGGGGAAAGTCATAACA
AACCATCAGGTTCTGGGAATGATAAAAGTGTACCACTTCAGCATCAGGAAGTGAAAATCTTACT
CTTATTCAACAGGAAGTGGATGCTTGGAAAGATAAGCAGGAGCTTCTGGAAACAGCTGATCT
ATATGCTACCAAGGAGAGAATAGAATACTCCAAACCTCAAGGGAAATATTTAATTTCTGGTT
ACTTTCTCTATTACTGTGTTGGAAATTTCATGGCTACCATCAATACTGTTTGATCGAGTT
GGGAAAACGGATCCTGTCAAAAGAGGCATTGAGATCACTGTGAATTATCTGGGAATCCAATTGATGT
GAAGTTTGGTCCAACACATTCTCATTCTGGAAATAATCGTCACATCCATCAGAGGAT
TGCTGATCACTCTACCAAGTTCTTATGCCATCTAGCAGTAAGTCCTCCAATGTCATTGTCCTG
CTATTAGCACAGATAATGGCATGTACTTGTCTCTGTGCTGATCCGAATGAGTATGCCCTT
AGAATACCGACCATAATCACTGAAGTCCTGGAGAACTGCAGTTCAACTCTATCACCGTTGGTTG
ATGTGATCTCCTGGTCAGCGCTCTCTAGCATACTCTCCTCTATTGGCTCACAAACAGGCACCA
GAGAAGCAAATGGCACCT**TGA**ACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTCAAATTAA
GATATAAGAGGGGGAAAATGGAACCAGGGCTGACATTATAAACAAACAAATGCTATGGTAGC
ATTTTCACCTTCATAGCATACTCCTCCCCGTAGGTGATACTATGACCATGAGTAGCATGCCAG
AACATGAGAGGGAGAACTAACTCAAGACAATACTCAGCAGAGGCATCCGTGGATATGAGGCTGG
TGTAGAGGCGGAGAGGCCAGAAACTAAAGGTAAAAACTGGAACCTGGAACTCTGGGGCAAGACATGT
CTATGGTAGCTGAGCCAAACACGTAGGATTCCGTTAAGGTTACATGGAAAAGGTTAGCTTGC
CCTTGAGATTGACTCATAAAATCAGAGACTGTAACAAAAAAAGGGCGGCCGCG
ACTCTAGAGTCGACCTGCAGAAGCTGGCCGCCATGGCCCAACTGTTATTGCAGCTATAATG

FIGURE 4

MSFLIDSSIMITSQLILFFFGFWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV
LNSSSRYFHWKMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFFWKLGD
FPILSPKHGILSIEQLISRVGIVGVTLMALLSGFGAVNCPTYMSYFLRNVTDILALERRLLQ
TMDMIISKKRMAMARRTMFQKGEVHNKPSGFWMGIKSVTTSASGSENLTLIQQEVDALEELSRO
LFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGKTDPVTRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLAQIMGY
FVSSVLLIRMSMPELEYRTIITEVLGELQNFYHRWFDVIFLVSALSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTCTCGTTATGAAGTGGAGCAGTGAGTGTGAGCCTAACATAGTTCC
AGAACTCTCCATCCGGACTAGTTATTGAGCATTGCCTCTCATATCACCAGTGGCCATCTGAGGT
GTTTCCCTGGCTCTGAAGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTGCTCTCACT
TCCATCTGGACCACGAGGCTCTGGTCCAAGGCTTTGCGTGAGAAGAGCTTCCATCCAGGT
GTCATGCAGAATTATGGGATCACCCTGTGAGCAAAAGGCGAACAGCAGCTGAATTTCACAG
AAGCTAAGGAGGCCTGTAGGCTGCTGGACTAAGTTGCCGGCAAGGACCAAGTTGAAACAGCC
TTGAAAGCTAGCTTGAACACTGCAGCTATGGCTGGGTTGGAGATGGATTGTCATCTCTAG
GATTAGCCAAACCCAAAGTGTGGAAAAATGGGTTGGGTGCTGATTGAAAGGTTCCAGTGA
GCCGACAGTTGCAGCCTATTGTTACAACACTCATCTGATACTTGGACTAACCTGTCATTCCAGAA
ATTATCACCAACCAAAAGATCCCATATTCAACACTCAAACGCAACACAAACAGAATTATTGT
CAGTGACAGTACCTACTCGGTGGCATCCCCTACTCTACAATACCTGCCCTACTACTCCTC
CTGCTCCAGCTTCACTTCTATTCCACGGAGAAAAAAATTGATTGTCACAGAAGTTTATG
GAAACTAGCACCATGTCTACAGAAACTGAACCAATTGTTGAAATAAAGCAGCATTCAAGAATGA
AGCTGCTGGGTTGGAGGTGCCCCACGGCTCTGCTAGTGCTTGCTCTCCTTTGGTGTG
CAGCTGGCTTGGATTGCTATGTAAAAGGTATGTGAAGGCCTCCCTTACAAACAAGAAT
CAGCAGAAGGAAATGATCGAAACCAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCTAA
TGAGGAATCAAAGAAAATGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAAACTACCGTGC
GATGCCTGGAAGCTGAAGTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTCTT
CATGCTCCTTACCCCTGCCCAAGCTGGGAAATCAAAGGCCAAAGAACCAAAGAACAGAATGCA
CCCTGGTTCTTAACGGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAAT
GCCCTCTCCTTATTGTAACCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTCCCACGGCT
TTCTAGCCTGGCTATGCTTAATAATATCCCACCTGGAGAAAGGAGTTGCAAAGTGCAGGAC
CTAAAACATCTCATCAGTATCCAGTGGTAAAAGGCCCTGGCTGTCTGAGGCTAGGTGGTTG
AAAGCCAAGGAGTCAGTGAGACCAAGGCTTCTACTGATTCCGAGCTCAGACCCCTTCTCA
GCTCTGAAAGAGAACACGTATCCCACCTGACATGTCCTCTGAGCCCGTAAGAGCAAAAGAAT
GGCAGAAAAGTTAGCCCCTGAAAGCCATGGAGATTCTCATAACTGAGACCTAATCTCTGAAA
GCTAAAATAAAGAAATAGAACAAAGGCTGAGGATACGACAGTACACTGTCAGCAGGACTGAAAC
ACAGACAGGGTCAAAGTGTCTCTGAACACATTGAGTTGAATCACTGTTAGAACACACACA
CTTACTTTCTGGTCTTACCACTGCTGATATTCTCTAGGAAATATACTTTACAAGTAACA
AAAATAAAACTCTTATAAATTCTATTGAGTTACAGAAATGATTACTAAGGAAGATT
ACTCAGTAATTGTTAAAAGTAATAAAATTCAACAAACATTGCTGAATAGCTACTATATGTC
AAGTGCCTGCAAGGTATTACACTCTGTAATTGAATATTATTCCCTAAAAATTGCACATAGTAG
AACGCTATCTGGGAAGCTATTGTTCTAGTTGATATTCTAGCTTATCTACTTCCAAACTAAT
TTTATTGCTGAGACTAATCTTATCATTCTCTAATATGGCAACCATTATAACCTTAATT
TATTATTAACATACCTAAGAAGTACATTGTTACCTCTATATACCAAGCACATTAAAAGTGC
ATTAACAAATGTACTAGCCCTCCTTCCAAACAAGAAGGGACTGAGAGATGCAGAAATATT
TGTGACAAAAATTAAAGCATTAGAAAACCTT

FIGURE 6

MARCFSLVLLLTSIWTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLG
LSLAGKDQVETALKASFETCSYGVGDGFVVISRISPNSPKCGKNGVGVLIWKVPVSRQFAAYCYN
SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTPPAPASTSIPR
RKKLICVTEVFMETSTMSTETEPFVENKAASFKNEAAGFGGVPTALLVLALLFFGAAAGLGFCYVK
RYVKAFFPTNKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCGCGCTCCCGACCCGGCCCGCCACCGCGCCGCTCCGCATCTGCACCCGAGCCCGC
GGCCTCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCGAGCGCAACTCGGTCCAGTCGGGGCG
CGGCTCGGGCGCAGAGCGGAGATGCAGCGGCTTGGGGCACCTGCTGTGCCTGCTGGCG
CGCGGTCCCCACGGCCCCCGCGCCGCTCCGACGGCAGCTCGGCTCCAGTCAAGCCGGCCG
GCTCTCAGCTACCCGCAGGAGGCCACCCCTCAATGAGATGTTCCGGAGGTTGAGGAACGTGAT
GGAGGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCTGCTGCTA
AAGCATCATCAGAAGTGAACCTGGCAAACCTACCTCCCAGCTATCACAATGAGACCAACACAGAC
ACGAAGGTTGGAAATAATACCATCCATGTGCACCGAGAAATTACAAGATAACCAACAACCAGAC
TGGACAAATGGTCTTTCAGAGACAGTTACATCTGTGGGAGACGAAGAAGGCAGAAGGAGGCC
ACGAGTGCATCATCGACGAGGACTGTGGGCCAGCATGTACTGCCAGTTGCCAGCTCCAGTAC
ACCTGCCAGCCATGCCGGGCCAGAGGATGCTCTGCACCCGGACAGTGAGTGCTGTGGAGACCA
GCTGTGTCTGGGTCACTGCACCAAAATGGCCACCAGGGCAGCAATGGGACCATCTGTGACA
ACCAGAGGACTGCCAGCCGGGCTGTGCTGTGCCTCAGAGAGGCCAGCTGTTCCCTGTG
ACACCCCTGCCCGTGGAGGGCGAGCTTGCATGACCCCGCCAGCCGGCTCTGGACCTCATCAC
CTGGGAGCTAGAGCCTGATGGAGCCCTGGACCGATGCCCTGTGCCAGTGGCCTCCCTGCCAGC
CCCACAGCCACAGCCTGGTGTATGTGTGCAAGCCGACCTCGTGGGAGCCGTGACCAAGATGG
GAGATCCTGCTGCCAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTCATGGAGGAGGTGCG
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGAGCCTGCCAG
CCGCCGCTGCACTGCTGGAGGGAAAGAGATTTAGATCTGGACCAGGCTGTGGTAGATGTGCAA
TAGAAATAGCTAATTATTCAGGTGTGCTTAGGCAGGTGGCTGACCAGGCTTCTTCTA
CATCTTCTTCCCAGTAAGTTCCCCTCTGGCTTGACAGCATGAGGTGTTGCAATTGTTAGCT
CCCCCAGGCTGTTCTCCAGGCTCACAGTCTGGTGTGGAGAGTCAGGCAGGGTAAACTGCA
GGAGCAGTTGCCACCCCTGTCCAGATTATTGGCTGCTTGCCTCTACCAAGTTGGCAGACAGCG
TTTGTCTACATGGCTTGATAATTGTTGAGGGAGGAGATGAAACAATGTGGAGTCTCCCTC
TGATTGGTTTGGGAAATGTGGAGAAGAGTGCCCTGCTTGCAAAACATCAACCTGGAAAAATG
CAACAAATGAATTTCACGCAGTTCTTCATGGCATAAGCTGTGCTTCAAGCTGCTCAGCTGC
AGATGAAATGTTCTGTTCACCCCTGCATTACATGTGTTATTCACTCCAGCAGTGTGCTCAGCTCC
TACCTCTGTGCCAGGGCAGCATTTCATATCCAAGATCAATTCCCTCTCAGCACAGCCTGGGG
AGGGGGTCATTGTTCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTGCC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTCATCTGGTGTGACTCTAAGCTCAGTGCT
CTCCACTACCCACACCAGCCTGGTGCCACCAAAAGTGTCTCCCAAAAGGAAGGAGAATGGGAT
TTTCTTGAGGCATGCACATCTGAAATTAGGTCAAACATAATTCTCACATCCCTCTAAAGTAAA
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGGCAGCCGTCTTCAATGAAGACAATGAT
ATTGACACTGTCCCTCTTGGCAGTTGCATTAGTAACTTGAAAGGTATATGACTGAGCGTAGCA
TACAGGTTAACCTGCAGAAACAGTACTTAGGTAAATTGTAGGGCGAGGATTATAATGAAATTG
AAAATCACTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGC
TGTGTGAAACATGGTTGTAATATGCGACTGCGAACACTGAACCTACGCCACTCCACAAATGATG
TTTCAGGTTGTCATGGACTGTTGCCACCATGTATTCACTCCAGAGTTCTAAAGTTAAAGTTGCA
CATGATTGTATAAGCATGCTTCTTGAGTTAAATTATGTATAAACATAAGTTGCATTAGAA
ATCAAGCATAAAATCACTCAACTGCAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQHKL
RSAVEEMEAEEAAKASSEVNLNLPPSYHNETNTDTKGNNTIHVHREIHKITNNQTGQMVFSE
TVITSVGDEEGRRSHECI IDEDCGPSMYCQFASFQYTCQPCRGQRLCTRSECCGDQLCVWGHC
TKMATRGSNGTICDNQRDCQPGLCCAFQRGLFPVCTPLPVEGELCHDPASRLLDLITWELEPDG
ALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFMEEVRQELEDLE
RSLTEEMALGEPAAAAALLGEEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGCGGACCGTGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCCAC
GGCCCACCTGTGAACTCCTCGTGCCAGGGCTGATGTGCGTCTCCAGGGCTACTCATCAAAG
GCCTAATCCAACGTTCTGTCTCAATCTGCAAATCTATGGGTCTGGGGCTTTCTGGACCCCTT
AACTGGGTACTGCCCTGGCCAATGCGTCCTCGCTGGAGCCTTGCCTCCTCTACTGGCCTT
CCACAAGCCCCAGGACATCCCTACCTCCCTTAATCTCGCCTCATCCGCACACTCCGTTACC
ACACTGGGTATTGGCATTGGAGCCCTCATCCTGACCCCTGTGCAGATAGCCCAGGTATTTG
GAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCTGTAGCCCGCTGCATCATGTGCTGTT
CAAGTGCTGCCTCTGGTGTCTGGAAAAATTATCAAGTCTAAACCGCAATGCATACATCATGA
TCGCCATCTACGGGAAGAATTCTGTCTCAGCCAAAATGCGTCATGCTACTCATGCGAAC
ATTGTCAGGGTGGTCGTCTGGACAAAGTCACAGACCTGCTGCTGTTCTTGGAGCTGCTGGT
GGTCGGAGGCGTGGGGCTCTGTCTTTCTCCGGTCGCATCCGGGGCTGGTAAAG
ACTTTAAGAGCCCCCACCTCAACTATTACTGGCTGCCATCATGACCTCCATCCTGGGGCCTAT
GTCATGCCAGCGGCTTCAGCGTTTCGGCATGTGTGGACACGCTTCCTGCTTCT
GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTCTAA
AGATTCTGGCAAGAAGAACGAGGCGCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGG
CCCTGATCCAGGACTGCACCCACCCACCGTCCAGCCATCCAACCTCACTCGCCTACAGGT
CTCCATTTGTGGTAAAAAAAGTTAGGCCAGGCGCGTGGCTACGCCTGTAATCCAACACT
TTGAGAGGCTGAGGCGGGCGGATCACCTGAGTCAGGAGTTGAGACCAGCCTGGCCAACATGGT
AAACCTCCGTCTATTAAAAATACAAAAATTAGCCAGAGTGGTGGCATGCACCTGTCATCCCA
GCTACTCGGGAGGCTGAGGCAGGAGAACGCTTGAACCCGGAGGCAGAGGTTGCAGTGAGCCGA
GATCGGCCACTGCACCTGGGTGACAGACTCTGTCTCCAAAACAAACAAACAAA
AAGATTTTATTAAAGATATTGTTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHNVSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLFWTL
NWVLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAFMLLMRN
IVRVVVLDKVTDLFFFKGKLLVGGVGVLFFFSGRIPGLGKDFKSPHLNYYWLPIMTSILGAY
VIASGFFSVFGMCVDTLFLCFILEDLERNNGSLDRPYMSKSLKILGKKNEAPPDNKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

CCCCCGCGCCCGGCGCCGGCGCCGAAGCCGGGAGCCACCGCCATGGGGCCTGCCTGGGAGCCTGC
TCCCTGCTCAGCTGCGCGTCTGCCTCTGCGGCTCTGCCCTGCATCCTGTGCAGCTGCTGCCCGC
CAGCCGCAACTCCACCGTGAGCCGCCTCATCTTACGTTCTCCTCTCCTGGGGTGCTGGTGTCCA
TCATTATGCTGAGCCCGGGCGTGGAGAGTCAGCTACAAGCTGCCCTGGGTGTGTGAGGAGGGGCC
GGGATCCCCACCGTCCCTGCAGGCCACATCGACTGTGGCTCCCTGCTGGCTACCGCCTGTACCG
CATGTGCTCGCCACGGCGCCTCTTCTTCTTACCCCTGCTCATGCTCTGCGTGANCGAGCA
GCCGGGACCCCCGGGCTGCCATCCAGAATGGGTTGGTCTTAAGTTCTGATCCTGGTGGCCTC
ACCGTGGGTGCCTCTACATCCCTGACGGCTCCTCACCAACATCTGGTTCTACTTCGGCGTGTGG
CTCCTCCTCTTCATCCTCATCCAGCTGGTGTGCTCATCGACTTGCACACTCCTGGAACCAGCGGT
GGCTGGCAAGGCCGAGGAGTGCATTCCCGTGCCTGGTACGCAGGCCCTTCTTACTCTCCTC
TTCTACTTGCTGTCGATCGCGGCCGTGGCGTGTGTTCATGTACTACACTGAGCCCAGCGGTGCCA
CGAGGGCAAGGTCTTCATCAGCCTAACCTCACCTCTGTGTCTGCGTGTCCATCGCTGCTGTCC
CCAAGGTCCAGGACGCCAACGCCAACCTGGGTCTGCTGCAGGCCCTGGTACCGCCTCACCC
TTTGTCACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCCATTGCCAACCC
GGGCAACGAGACAGTTGTGGCAGGCCCGAGGGCTATGAGACCCAGTGGTGGATGCCCGAGCATTG
TGGGCCTCATCATCTCCTCTGTGCACCCCTTCATCAGTCTGCCTCCTCAGACCACCGCAGGTG
AACAGCCTGATGCAGACCGAGGAGTGCCACCTATGCTAGACGCCACACAGCAGCAGCAGCAGGT
GGCAGCCTGTGAGGCCGGGCTTGTACAACGAGCAGGACGGCGTACCTACAGCTACTCCTCTCC
ACTTCTGCCTGGTGTGGCTACTGCACGTATGACGCTACCAACTGGTACAAGCCGGTGAG
ACCCGGAAGATGATCAGCACGTGGACCGCCGTGTGGTGAAAGATCTGTGCCAGCTGGCAGGGCTGCT
CCTCTACCTGTGGACCCCTGGTAGCCCCACTCCTCTGCACCGCAGCTCAGGTGAGGCAGCCTCA
CAGCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCTCTCGGCTCGGTGACAGCCAACCTGCC
CCCCACACCAATCAGCCAGGCTGAGCCCCACCCCTGCCAGCTCCAGGACCTGCCCTGAGCCGGC
CTTCTAGTCGTAGTGCCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCAGAGCCCCATCCCCCG
ACCCACACGGTGGAGCTGCCTCTCCTCCCTCCCTGTTGCCATACTCAGCATCTGGATGAA
AGGGCTCCCTGTGCTCAGGCTCCACGGGAGCGGGCTGCTGGAGAGAGCGGGGAACCTCC
TGGGGCATCCGGCACTGAAGCCCTGGTGTGCTGGTACGTCCCCCAGGGACCCCTGCCCT
GACTCGTGCCTACTGAGTCTAAGACTTTCTAATAAACAGCCAGTGCCTGTAAAAAAA

FIGURE 12

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVESQL
YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFTLLMLCVSSSRDPRAAIQ
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSLFILQLVLLIDFAHWNQRWLGKAE
ECDSRAWYAGLFFFLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFCVCVSIAAVLPKV
QDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI
VGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQQVAACEGRAFDNEQDGVTYSY
SFFHFCLVLASLHVMMTLTNWYKPGTRKMISTWTAVWVKICASWAGLLLWTLVAPLLLNRD
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGGCCAGGAACCACCGTTAAGGTCTTCTCTTTAGGGATGGTGA
GGTTGGAAAAAGACTCCTGTAACCCTCCTCCAGGATGAACCACCTGCCAGAAGACATGGAGAACG
CTCTCACCGGGAGCCAGAGCTCCATGCTCTCGCAGAACATCCATTCCATCAACCCCCACACAA
CTCATGCCAGGATTGAGTCCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC
TTTCTGTTGTTGTCACCTTGACCTCTTATTGTAACATTACTGTGGATAATAGAGTTAAATG
TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTATTCTCATAT
TTTGATATATTCTTCTGGCAGTTTCGATTAAAGTGTAAACTTGATATGCTGTGCAG
ACTGCGCCATTGGTGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTACTAGCAAAG
TGATCCTTCGAAGCTTCTCAAGGGCTTGGCTATGTGCTGCCATCATTCAATC
CTTGCCTGGATTGAGACGTGGTCCTGGATTCAAAGTGTACCTCAAGAACAGAAGAAA
CAGACTCCTGATAGTCAGGATGCTCAGAGAGGGCAGCACTTACCTGGTGGCTTCTGATG
GTCAGTTTATTCCCTCCTGAATCCGAAGCAGGATCTGAAGAACGCTGAAGAAAAACAGGACAGT
GAGAAACCACTTAGAACTATGAGTACTACTTTGTTAAATGTGAAAAACCCCTCACAGAAAGTC
ATCGAGGCACAAAGAGGCAGGCAGTGAGTCTCCCTGTCGACAGTAAAGTGAATGGTACGTC
CACTGCTGGTTATTGAACAGCTAATAAGATTATTGTAATACCTCACAAACGTTGTAC
CATATCCATGCACATTAGTGCCTGCCTGTTGGCTGGTAAGGTAATGTCATGATTCACTCT
TCAGTGAGACTGAGCCTGATGTGTTAACAAATAGGTGAAGAAAGTCTGTGCTGTATTCTAAC
AAAAGACTTAATATTGAAGTAACACTTTAGTAAGCAAGATACTTTATTCAATTAC
AGAATGGAATTTTTGTTCATGTCTCAGATTATTTGTTCTTTAACACTCTACATT
TCCCTGTTTTAACTCATGCACATGTGCTCTTGTACAGTTAAAAAGTGTAAATAAAATCTG
ACATGTCAATGTGGCTAGTTTATTCTGTTGCATTATGTGTATGGCCTGAAGTGTGGA
CTTGCAAAAGGGAGAAAGGAATTGCGAATACATGTAAAATGTCACCAGACATTGTATTATT
TTATCATGAAATCATGTTCTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTGAAATGC
ACAAAATGACTAAACCATTATCATGTTCCCTTGCCTCAGCCAATTCAATTAAAATGAA
CTAAATTAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDLLF
VTLLWIIELNVNGGIENTLEKEVMQYDYYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALT
AVTSAFLLAKVILSKLFSQGAFGYVLPIISFILAWIETWFLDFKVLPOEAEEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSEEAEKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTTCGGACCCAGGACCCCTCGGGCCGACCCGCCAGGAAAGACTGAGG
CCGCGGCCTGCCCCGCCGGCTCCCTGCGCCGCCGCTCCCAGGACAGAAAGATGTGCTCCAG
GGTCCCTCTGCTGCTGCCGCTGCTCTGCTACTGGCCCTGGGGCTGGGGTGCAGGGCTGCCCAT
CCGGCTGCCAGTGCAGCCAGCACAGACAGTCTCTGCACTGCCGCCAGGGGACCAAGGTGCC
CGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTGAGAACGGCATCACCATGCTCGACGC
AGGCAGCTTGCCGGCTGCCGGCTGCAGCTCTGGACCTGTACAGAACAGATGCCAGGC
TGCCCAGCGGGTCTTCAGCCACTGCCAACCTCAGCAACCTGGACCTGACGGCCAACAGGCTG
CATGAAATACCAATGAGACCTCCGTGGCCTGCGGCCCTGAGCGCCTCTACCTGGCAAGAA
CCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGCTCGACCGCCTCTGGAGCTCAAGCTGC
AGGACAACGAGCTGCCGGACTGCCCGCTGCCCTGCCCTGCTGCTGGACCTCAGC
CACAAACAGCCTCTGCCCTGGAGGCCGATCTGGACACTGCCAACGTGGAGGCCTGCC
GGCTGGTCTGGGGCTGCAGCAGCTGGACGAGGGCTTTCAGCCGCTGCGAACCTCCACGACC
TGGATGTGTCGACAACCAGCTGGAGCGAGTGCACCTGTGATCCGAGGCCCTGGGGCTGACG
CGCCTGCGGCTGCCGGCAACACCCGATTGCCAGCTGCCGCCAGGACCTGCCGGCCTGGC
TGCCCTGCAAGGAGCTGGATGTGAGCAACCTAACGCTGAGGCCCTGCGCTGGGACCTCTCGGGC
TCTTCCCCCGCTGCCGTGCTGGCAGCTGCCCAACCCCTCAACTGCGTGTGCCCTGAGC
TGGTTGGCCCTGGGTGCGAGAGGCCACGTACACTGCCAGCCCTGAGGAGACGCGCTGCCA
CTTCCGCCCAAGAACGCTGCCGGCTGCTCTGGAGCTGACTACGCCACTTGGCTGCCAG
CCACCACCACACAGGCCACAGTGCACCGAGGCCCTGCGTGGGAGGCCACAGCCTGTCT
TCTAGCTTGGCTCTACCTGGCTAGCCCCACAGGCCGGCCACTGAGGCCCTAGCCCGCC
CACTGCCCAACCGACTGTAGGCCCTGTCCTCCAGGCCACTGCCACCGTCCACCTGCC
ATGGGGGACATGCCACCTGGGGACACGGCACCCACTGGCGTGTGCCCCGAAGGCTTCAGC
GGCCTGTACTGTGAGAGGCCAGATGGGGCAGGGACACGCCCTACACCAGTCACGCCAG
GCCACCACGGTCCCTGACCCCTGGCATCGAGCCGGTGGAGGCCACCTCCCTGCGTGGGGCTGC
AGCGCTACCTCCAGGGAGCTCCGTGCACTCAGGAGCCCTCGTCACTATCGAACCTATCG
GGCCCTGATAAGCGGTGGTGACGCTGCGACTGCCCTGCCCTCGCTGAGTACACGGTACCCCA
GCTGCCGCCAACGCCACTTACTCCGTGTGTCATGCCCTTGGGCCGGGCGGGTGCAGG
GCGAGGAGGCCCTGCCGGAGGCCATACACCCCGGCCCTGCGTCACTCCAAACCACGCC
CAGGCCCGCAGGGCAACCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCC
GCTGGCTGCCGTGGGGCAGCTACTGTGTCGGCGGGGGCGGGCATGGCAGCAGGGCTCAGG
ACAAAGGGCAGGTGGGGCAGGGCTGGGCCCTGGAACCTGGAGGGAGTGAAGGTCCCC
CCAGGCCCGAAGGCCACAGAGGGCGGTGGAGAGGCCCTGCCAGCGGGTCTGAGTGTGAGG
ACTCATGGCTTCCCAGGGCTGGCTCCAGTCACCCCTCCAGC
GAGAGAGACAGGGCAGCTGGGCCGGCTCAGCCAGTGA
ACACCACGTAAGTCTCAGTCCAACCTGGGGATGTGTCAGACAGGGCTGTGAC
GGGCCCTGTTCCCTGGACCTCGGTCTCCTCATCTGTGAGATGCTGTGGGCCAGCTGAC
CTAACGTCCCCAGAACGAGTGCCTATGAGGACAGTGTCCGCC
CCTGGGCACGGCGGGCCCTGCCATGTGCTGGTAACGCATGCC
TCCAGGCAGGCCCTGGGGCCAGTGAAGGAAGCTCCGGAAAGAGCAGAGGGAGAGCGGG
GGCTGTGTGACTCTAGTCTTGGCCCGAGGAAGGAACAAAGAAACTGGAAAGGAAGATGC
TTTAGGAACATGTTGCTTTAAATATATATTTATAAGAGATC
GGGAAGATTTTCAAACCTCAGAGACAAGGACTTGGTTTTGTAAGAC
GGCTTTGTAAGAAAAATAAAAGATGAAGTGTGAAA

FIGURE 16

MCSRVPLLLPLLLALGPGVQGCPGQCQCSQPQTVFCTARQGTTVPRDVPPDTVGLYVFENGIT
MLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRLERLY
LGKNRIRHIQPGAFDTLDRLLELKLDNELRALPPLRLPRLLLDLSHNSLLALEPGILDGTANVE
ALRLAGLGLQQLDEGLFSRLRNLHDLDVSDNQLERVPPVIRGLRGLTRLRLAGNTRIAQLRPEDL
AGLAALQELDVSNLSLQALPGDLSGLFPRLRLLAAARNPNCVCPLSWFGPWVRESHVTLASPEE
TRCHFPPKNAGRLLLELDYADFGCPATTTATVPTTRPVREPTALSSSLAPTWLSPTAPATEAP
SPPSTAPPTVGPVPQPQDCPPSTCLNGGTCHLGTRHHLACLCPEGFTGLYCESQMGQGTRPSPTP
VTPRPPRSLTLGIEPVSPSLRVGLQRYLQGSSVQLRSLRLTYRNLSGPDKRLVTLRLPASLAEY
TVTQLRPNATYSVCVMPILGPGRVPEGEAACGEAHTPPAVHSNHAPVTQAREGNLPLLIAPALAAV
LLAALAAVGAAYCVRRGRAMAAAQDKGQVGPAGPLELEGVKVPLEPGPKATEGGEALPSGSE
CEVPLMGFPGPGLQSPLHAKPYI

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 17

GCAGCGCGAGGCAGCGGTGGCTGAGTCGTGGCAGAGCGAAGCGACAGCTCATGCG
GGTCGGATAGGGCTGACGCTGCTGTGCGGTGCTGCTGAGCTTGGCCTCGCGTCTCGG
ATGAAGAAGGCAGCCAGGATGAATCCTAGATTCCAAGACTACTTGACATCAGATGAGTCAGTA
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTCTTGTGATTCAAAGAATCTGA
ATTAGAATCCTCTATTCAAGAAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGAAAGTGTACAG
AAGATATCAGCTTCTAGAGTCTCAAATCCAGAAAACAAGGACTATGAAGAGCCAAGAAAGTA
CGGAAACCAGCTTGACCGCCATTGAAGGCACAGCACATGGGAGCCCTGCCACTTCCCTTTCT
TTTCCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAAGATGGCAGACTGTGGTGTG
CTACAACCTATGACTACAAAGCAGATGAAAAGTGGGCTTTGTGAAACTGAAGAAGAGGCTGCT
AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAACGTGAATGAAAATCTTAATGGAAG
CAATAAGAAAAGCCAAAAAAGAGAAGCATATCGGTATCTCCAAAAGGCAGCAAGCATGAACCATA
CCAAAGCCCTGGAGAGAGTGTATGCTCTTATTGGTGTGATTACTGCCACAGAATATCCAG
GCAGCGAGAGAGATTTGAGAAGCTGACTGAGGAAGGCTCTCCCAAGGGACAGACTGCTCTGG
CTTCTGTATGCCCTGGACTTGGTTAATTCAAGTCAGGCAAAGGCTTGTATATTACAT
TTGGAGCTCTGGGGCAATCTAATAGCCCACATGGTTGGTAAGTAGACTTTTAGTGGAAAGGCT
AATAATATTAACATCAGAAGAATTGGTTATAGCGGCCACAACCTTTCAGCTTCATGATC
CAGATTGCTGTATTAAGACCAAATTCAGTGAACTCCTCAAATTCTTGTAAATGGATAT
AACACATGGAATCTACATGAAATGAAAGTTGGTGGAGTCCACAATTTCCTTAAATGATTAG
TTTGGCTGATTGCCCTAAAAGAGAGATCTGATAAATGGCTTTAAATTTCTCTGAGTTG
GAATTGTCAGAATCATTTCACATTAGATTATCATAATTAAAAATTTCCTTAGTTCA
AAATTGTAAATGGGGCTATAGAAAAACACATGAAATATTACAATATTGCAACAATGC
CCTAAGAATTGTTAAATTGAGTTATTGTCAGAATGACTCCAGAGAGCTCTACTTCTG
TTTTTACTTTCATGATTGGCTGTCTCCCATTATTCTGGTCAATTGCTAGTGACACTGT
GCCTGCTCCAGTAGTCTCATTCCCTATTGCTAATTGTTACTTTCTTGCTAATTGG
AAGATTAACTCATTAAATAAAATTATGCTAAGATTTTTTTTTTTTTTTTTTTTTTT
AA

FIGURE 18

MRVRIGLTLLLCAVLLSLASASSDEEGSQDESLDSTTTSDESVKDHTTAGRVVAGQIFLDSEESEL
ESSIQEEEDSLKSQEGERVTEDISFLESPNPNKDYEPPKKVRKPALTAEIGTAHGEPCFPFLFLDK
EYDECTS DGDREDGRLWCATTYDYKADEKWGFCETEEAAKRRQMQEAEEMMYQTGMKILNGSNKKSQKR
EAYRYLQKAASMNHTKALERVSYALLFGDYLPPQNIQAAREMFEKLTEEGSPKGQTALGFLYASGLGVN
SSQAKALVYYTFFGALGGNLIAHMVLVSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCA GATT TAAG CCCATTCTGCAGT GGAATT CATGAA CTAGCAAGAGGACACC ATCTTCTT
GTATTATA CAAGAAAGGAGTGTACCTATCACACACAGGGGGAAAATGCTCTTGGGTGCTAGG
CCTCCTAATCCTCTGTGGTTCTGTGGACTCGTAAAGGAAA ACTAAAGATTGAAGACATCACTG
ATAAGTACATTTTATCACTGGATGTGACTCGGGCTTGGAAACTGGCAGCCAGAACTTTGAT
AAAAAGGGATTTCATGTAATCGCTGCCGTCTGACTGAATCAGGATCAACAGCTTAAAGGCAGA
AACCTCAGAGAGACTTCGTACTGTGCTCTGGATGTGACCGACCCAGAGAAATGTCAAGAGGACTG
CCCAGTGGGTGAAGAACCAAGTTGGGAGAAAGGTCTCTGGGTCTGATCAATAATGCTGGTGT
CCCGCGTGTGGCTCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA
CCTGTTGGACTCATCAGTGTGACACTAAATATGCTCCCTTGGTCAAGAAAGCTCAAGGGAGAG
TTATTAAATGTCTCCAGTGTGAGGTGCCTGCAATCGTTGGAGGGGCTATACTCCATCCAAA
TATGCAGT GGAAGGTTCAATGACAGCTTAAGACGGGACATGAAAGCTTTGGTGTGCACGTCTC
ATGCATTGAACCAGGATTGTTCAAAACAAACTGGCAGATCCAGTAAAGGTAAATTGAAAAAAAC
TCGCCATTGGGAGCAGCTGTCTCCAGACATCAAACAACATGGAGAAGGTTACATTGAAAAA
AGTCTAGACAAACTGAAAGGCAATAATCCTATGTGAACATGGACCTCTCCGGTAGAGTG
CATGGACCACGCTCTAACAGTCTCCCTAAGACTCATTATGCCGCTGGAAAAGATGCCAAA
TTTCTGGATACCTCTGTCTCACATGCCAGCAGCTTGCAAGACTTTTATTGTTGAAACAGAAA
GCAGAGCTGGCTAATCCAAGGCAGTGTGACTCAGCTAACCAAAATGTCTCCAGGCTATGA
AATTGGCCGATTTCAAGAACACATCTCCTTTCAACCCATTCTTATCTGCTCCAACCTGGACT
CATTAGATCGTGTCTATTGGATTGCAAAAGGGAGTCCCACCATCGCTGGTAGGTATCCCAGGGT
CCCTGCTCAAGTTCTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCCTGCCCT
GTATTAGGCTTGCCTGCTGGTGTGATGTAAGGAAATTGAAAGACTTGCCATTCAAATGA
TCTTACCGTGGCCTGCCCATGCTTATGGTCCCCAGCATTACAGTAACCTGTGAATGTTAAGT
ATCATCTCTTATCTAAATATTAAAAGATAAGTCAACCCAAAAAA
AAAAAAAAAAAAA

FIGURE 20

MLFWVLGLLILCGFLWTRKGKLKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLTESG
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINNAGVPGVLAPTDWLTLEDY
REPIEVNLFGLISVTLNMLPLVKKAQGRVINSSVGGRLAIVGGGYTPSKYAVEGFNDSLRRDMK
AFGVHVSCIEPGLFKTNLADPVKVIEKKLAIWEQLSPDIKQQYGEGYIEKSLDKLGNKSYVNMD
LSPVVECMDHALTSLFPKTHYAAAGKDAKIFWIPLSHMPAALQDFLLLKQKAELANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGCGGCGGTAGCATGGAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTGTGCTCGGCG
CACTCGCTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTCTTCTGGGAAGTAAAA
GGTGAAGCCAAGAACAGCATTACTGATCCAAATGGATGATGTTGAAGTTGTTATAATTGA
CATTAGAAATATATTCCATGCTATCAGCTTTAGCTTTATAATTCTCAGGCGAAGTAAATG
AGCAAGCACTGAAGAAAATTATCAAATGTCAAAAAGAATGTGGTAGGTTGGTACAAATTCCGT
CGTCATTCAAGATCAGATCATGACGTTAGAGAGAGGCTGCTCACAAAAACTGCAGGAGCATT
TTCAAACCAAGACCTGTTCTGCTATTAACACCAAGTATAATAACAGAAAGCTGCTACTC
ATCGACTGGAACATTCTTATATAAACCTCAAAAAGGACTTTCACAGGGTACCTTAGTGGTT
GCCAATCTGGGCATGCTGAACAACTGGTTATAAAACTGTATCAGGTTCTGTATGTCCACTGG
TTTAGCCGAGCAGTACAAACACACAGCTCTAAATTGGTAAAGAAGATGGATCCTAAAGGAGG
TACATAAGATAAAATGAAATGTATGCTTCATTACAAGAGGAATTAAAGAGTATATGCAAAAAGTG
GAAGACAGTGAACAAGCAGTAGATAACTAGTAAAGGATGTAAACAGATTAAACGAGAAATTGA
GAAAAGGAGAGGGAGCACAGATTCAAGGAGCAAGAGAGAAGAACATCCAAAAGACCCTCAGGAGA
ACATTTCCTTGTCAAGCATTACGGACCTTTCCAATTCTGAATTCTTCATTCACTGTGTT
ATGTCTTAAAAAATAGACATGTTCTAAAGTAGCTGTAACTACAACCACATCTCGATGTAGT
AGACAATCTGACCTTAATGGTAGAACACACTGACATTCTGAAGCTAGTCCAGCTAGTACACCAC
AAATCATTAAAGCATAAAGCCTAGACTAGATGACAGATGGCAATTCAAGAGATCTGGTTGTTA
GATACACAAGACAAACGATCTAAAGCAAATCTGGTAGTAGTAACCAAGATAAGCATCCAAAAT
GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTGGTGAATATTACGGTCTC
CTACATTTGATCCTTTAACCTACAAGGAGATTCTTATTGGCTGATGGTAAAGCCAAC
ATTCTATTGTTTACTATGTTGAGCTACTGCACTTACAAAGTACTTTCAAACATCA
TGTTGAGCTAATACACAGATAACTCTTAGTGCATTACAAAGTACTTTCAAACATCA
GATGCTTTATTCCAACCTTTTACCTTCACTAAGTTGAGGGAAAGGCTACACAG
ACACATTCTTAGAATTGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCCAGCACT
TAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGCAACGTATT
GAGACCATGTCTATTAAAAAATGGAAAAGCAAGAATAGCCTTATTTCAAAATATGGAAA
GAAATTATGAAAATTATGAGTCATTAAATTCTCCTTAAGTGTACTTTTAGAAGTA
CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAATTGCAAAACATCATCT
AAAATTAAAAAAAAAAAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTDSDEGFLGEVKGEAKNSITDSQMDDVEVVTIDIQKYI
PCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEHFSNQDL
VFLLLTPSIIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVGSCMSTGFSRAV
QTHSSKFFeedGSLKEVHKINEMYASLQEELKSICKVEDSEQAVDKLVKDVNRLKREIEKRRGA
QIQAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHSKSSCNYNHLDVVVDNLTL
MVEHTDIPEASPAStPQIIKHKALLDLDRWQFKRSRLLDTQDKRSKANTGSSNQDKASKMSSPET
DEEIEKMKGFGEYSRSPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGGGAGGGCAGAGTCAGCCGAGCCAGTCCAGCCGACGAGGGACAGCGCAGGGCAGCCCAA
GCAGCGCGAGCAACGCCGCCACACCCCTCTGCGGTCCCCGGCGCCTGCCACCCCTCCCTCCCC
GGTCCCCGCCCTGCCGGCCAGTCAGCTTGCGGGTTCGCTGCCCGGAAACCCGAGGTACCCAGCCCGCCTCT
GCTTCCCTGGGCCGCGCCCTCCACGCCCTCCTCTCCCTGGCCCGCCTGGCACCGGGACCGTTGCCCTGA
CGCGAGGCCAGCTACTTTGCCCGCGTCTCCGCCCTGCTGCCCTTCCACCAACTCCAACCTCCCTCTCCC
TCCAGCTCCACTCGCTAGTCCCGACTCCGCCAGCCCTCGGCCGCTGCCGTAGGCCGCTTCCGTCCGGTCCCAA
GGTGGGAACCGCGTCCGCCCGGCCACCATGGCACGGTCGCGTCCGCCGCTCTGCCACCCGGCAGTGCTC
AGGCCCGCGCTGCGTGGCTGCCAGCTCAAGTCGAAAAGTGTGCTCGGAAGTGCAGCTTTACGTGTCCAAAGGCTC
AACAGAACGATGCCCTCCACGAGATCAACGGTGTATTTGAAGATCTGTCCCCAGGGTCTACCTGCTGCTC
CAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGATGATTCAAAGTGTGGTCAGCGAACAGTGCAATCATTG
CAAGCTGTCTTGCTTACGTTACAAGAAGTTGATGAATTCTCAAAGAAACTACTTGAAAATGCAGAGAAATCCCTG
AATGATATGTTGTGAAGACATATGCCATTACATGAAAATTCTGAGCTATTAAAGATCTTCGTAGAGTTG
AACGTTACTACGTGGTGGAAATGTGAACCTGGAAGAAATGCTAAATGACTTCTGGCTGCCCTGGAGCGGATG
TTCCGCCTGGTGAACCTCCAGTACCACTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAG
CCCTTGGAGATGCTCCCTGCAAATTGAAGCTCCAGGTTACTCGTGTGTTGTAGCAGCCGTACTTCGCTCAAGGC
TTAGCGGTTGCGGGAGATGTCGTGAGCAAGGTCTCGTGGTAAACCCACAGCCAGTGTACCCATGCCCTGTTGAAG
ATGATCTACTGCTCCACTGCCGGGTCTCGTACTGTGAAGCCATGTTACAACACTGCTCAAACATCATGAGAGGC
TGTGTTGCCAACCAAGGGATCTGATTTGAATGAAACAATTCTAGATGCTATGCTGATGGTGGCAGAGAGGCTA
GAGGGTCTTCAACATTGAATCGGTATGGATCCCACGATGTGAAGATTCTGATGCTATTATGAACATGCAGGAT
AATAGTGTCAAGTGTCTCAGAAGGTTCCAGGGATGTGGACCCCCCAAGCCCCCTCCAGCTGGACGAATTCTCGT
TCCATCTGAAAGTGCCTCAGTGTGCTCAGACACATCACCCGAGGAACGCCAACACAGCAGCTGGCACT
AGTTGGACCGACTGGTACTGATGTCAAGGAGAAACTGAAACAGGCCAGAAATTCTGGTCTCCCTCCAGCAAC
GTTTGCACGATGAGAGGATGGCTGCAGGAAACGGAATGAGGATGACTGTTGGATGGAAAGGCAAAAGCAGGTAC
CTGTTGCAAGTGCAGGAAATGGATTAGCCAACCAGGGCAACAACCCAGAGGTCCAGGTTGACACCAGCAAACAGAC
ATACTGATCCTCGTCAAATCATGGCTTCTCGAGTGATGACCAGCAAGATGAAGAATGCATAACATGGAACGACGTG
GACTTCTTGATATCAGTGATGAAAGTAGTGGAGAAGGAAGTGGAGTGGCTGTGAGTATCAGCAGTGCCTTCAGAG
TTTGACTACAATGCCACTGACCAGTGTGGAGAGGTGCAATGAGAAAGCCGACAGTGCTGGTCCCTGGCAGGG
CAGGCCACCTCCACTGTCTCGCATCTGTTCTGGTATGCAGAGAGAGTGGAGATAATTCTCAAACACTGAG
AAAAAGTGTCAACCAAGTAAAGGCACCAAGTTACCTTCTACCCATCCTAGTGACTTGTGTTTTAAATGAA
TGGACAACAATGTACAGTTTACTATGTGGCACTGGTTAAGAAGTGCTGACTTGTGTTCTCATTGAGTTGGG
AGGAAAAGGGACTGTGCATTGAGTTGGTCTGCTCCCCAACCATGTTAAACGTGGCTAACAGTGAGGTACAGAA
CTATAGTTAGTTGTGCATTGTGATTATCACTCTATTATGTTGTATGTTTTCTCATTGAGTTGGG
TTTTTTCCAACGTGATCTGCCCTGTTCTACAGCAACAGGGCTTCTGGCACGTAACATGTACGTATT
TCTGAAATATTAATAGCTGTACAGAAGCAGGTTTATTTATCATGTTATCTTAAAGAAAAAGCCAAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRLYVSKGFNNDAPLHEINGDHLKICPQGST
CCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDFVKTYGH
LYMQNSELFKDLFVELKRYYVGVNVNLEEMLNDFWARLLERMFRVLVNSQYHFTDEYLECVSKYTE
QLKPFGDVPRKLKLQVTRAFVAARTFAQGLAVAGDVSKVSVNPTAQCTHALLKMIYCSHCRLG
VTVKPCYNCSNIMRGCLANQGDLDFEWNNFIDAMLMVAERLEGPFNIESVMDPIDVKISDAIMN
MQDNSVQVSQKVFOQCGPPKPLPAGRISRSISESAFSARFRPHPEERPTTAAGTSLDRLVTDVK
EKLKQAKKFWSLPSNCNDERMAAGNGNEDDCWNGKGKSRYLFAVTGNGLANQGNNEVQVDT
KPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAG
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC
CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTCATTATATTCTTCAAGCAACT
TACAGCTGCACCGACAGTTGCGATGAAAGTTCTAATCTCTCCCTCCTGTTGCTGCCACTAA
TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGTCGCCAGAGGCCACAGGGACCGAGGC
CAGGCTTCTAGGAGATGGCTCCAGGAAGGCGCCAAGAATGTGAGTGCAAAGATTGGTTCTGAG
AGCCCCGAGAAGAAAATTCATGACAGTGTCTGGGCTGCCAAAGAACAGCAGTGCCCTGTGATCATT
TCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGAAAGCCAACAAGCATTCCAGA
GCCTGCCAGCAATTCTCAAACAATGTCAGCTAAGAAGCTTGCTCTGCCTTGTAGGAGCTCTG
AGCGCCCACCTTCCAATTAAACATTCTCAGCCAAGAACAGCAGTGAGCACACCTACCAGACACTC
TTCTTCTCCCACCTCACTCTCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCGTCAGTCTAGCCT
TGTTTTCAAGATCATTGTTGCTCTCTAGTGTCTTCTCTCGTCAGTCTAGCCT
GTGCCCTCCCTTACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAAACTGTAGCTTCCCT
AGCTAGTGTCAATTAAACCTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAATATTTT
AAATGTCAAAAAAAAAAAAAAA

FIGURE 26

MKVLISLLLLPLMIMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM
TVSGLPKKQCPKCDHFKGKGVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAGCCAGTCCCCAGCGGAAGCACAGCTCAG
AGCTGGTCTGCCATGGACATCCTGGTCCCACTCCTGCAGCTGCTGGTGTGCTTACCTGCC
CCTGCACCTCATGGCTCTGCTGGCTGCTGGCAGCCCCTGTGCAAAGCTACTTCCCTACCTGA
TGGCCGTGCTGACTCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAGCTTCAGCCAG
ATAAAGGGCTTACAGGAGCCTCCGGAAAGTGGCCTACTGGAGCTGGCTGCCAACCGGAGC
CAACTTTCAGTTCTACCCACCAGGCTGCAGGGTACCTGCCTAGACCCAAATCCCCACTTGAGA
AGTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGTTGTGGCTGCCT
GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGTGCACCTGGTGTG
CTCTGTGCAGAGCCCAGGAAGGTCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTGC
TCTTTCTGGAGCATGTGGCAGAACCATATGGAAGCTGGCCTCATGTGGCAGCAAGTTTC
GAGCCCACCTGGAAACACATTGGGATGGCTGCCTCACCAAGAGACCTGGAAGGATCTTGA
GAACGCCAGTTCTCGAAATCAAATGGAACGACAGCCCCCTCCCTGAAGTGGTACCTGTTG
GGCCCCACATCATGGGAAAGGCTGTCAAACAATCTTCCAAGCTCCAAGGCACACTTTGCTCC
TTCCCCAGCCTCCAATTAGAACAGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACTA
GCAGAACATGAGAGAACATTCATGTACCCACCTACTAGTCCCTCTCCCCAACCTCTGCCAGGGC
AATCTCTAACTCAATCCGCCTCGACAGTGAAGAAAGCTCTACTTCTACGCTGACCCAGGGAGG
AAACACTAGGACCCCTGTTGTATCCTCAACTGCAAGTTCTGGACTAGTCTCCCAACGTTGCCTC
CCAATGTTGCCCTTCCTCGTCCATGGTAAAGCTCCTCTCGCTTCCCTGAGGCTACAC
CCATGCGTCTCTAGGAACCTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAACCCCCCTGAC
CCTCTCTCCCCACTACCACCTCTTGAGCTGGGGCACCAGGGAGAACATCAGAGATGCTGGGG
ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTGTTCTCAAATATTTTAATAAGACGAA
ACCACG

FIGURE 28

MDILVPLLQLLVLLLTLPLHLMALLGCWQPLCKSYFPYIMAVLTPKSNRKMESKKRELFSQIKGL
TGASGKVALLELGCCTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDVVVCTLVLCVQSPRKVLQEVRVLRPGGVLFWEHVAEPYGSWAFMWQQVFEPTW
KHIGDGCLTRETWKDLENAQFSEIQMERQPPPLKWLPGPHIMGKAVKQSFPSKALICSFPSL
QLEQATHQPIYLPLRGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATGTTGCCTATCCACCTCCCCAAGCCCCTTACCTATGCTGCTGCTAACGCTGCTGCTGCT
GCTGCTGCTGCTAAAGGCTCATGCTTGGAGTGGGACTGGTCGGTCCCAGAAAGTCTCTTCTG
CCACTGACGCCCATCAGGGATTGGCCTTCTTCCCCCTCCTTCTGTGTCTCCTGCCTCAT
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGAAGGGAGAAAGTGGGGATGGCTA
AGAAAGCTGGGAGATAGGAACAGAAGAGGGTAGTGGGTGGCTAGGGGGCTGCCTTATTAAA
GTGGTTTTATGATTCTTATACTAATTATAACAAAGATATTAAGGCCCTGTTCATTAAGAAATT
GTTCCCTTCCCTGTGTTCAATGTTGTAAGATTGTTCTGTGTAATATGTCTTATAATAAAC
AGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

MLLLTLLLLLKGSCLEWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTC SQAQPRG
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTCGAAATTCCCTCAACTATAACCCACAGTCAAAAGCAGACTCACTGTGTCCAGGCTACCAGTT
CCTCCAAGCAAGTCATTTCCCTTATTAAACCGATGTGTCCCTCAAACACACCTGAGTGCTACTCCCT
ATTGCATCTGTTGATAAATGATGTTGACACCCCTCCACCGAATTCTAAGTGAATCATGTCGG
GAAGAGATAACAATCCTGGCCTGTGTATCCTCGCATTAGCCTGTCTTGCCATGATGTTACC
TTCAGATTCATCACCAACCCCTCTGGTTCACATTTCATTGGTTATTGGGATTGTTGTT
TGTCTGCGGTGTTATGGTGGCTGTATTATGACTATAACCAACGACCTCAGCATAGAATTGGACA
CAGAAAGGAAAATATGAAGTGCCTGCTGGGTTGCTATCGTATCCACAGGCATCACGGCAGTG
CTGCTCGTCTGATTGGTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTCCAAATCAC
AAATAAGCCATCAGCAGTGCTCCCTCCTGCTGTTCCAGCCACTGTGGACATTGCCATCCTCA
TTTCTCTGGTCCTCTGGTGGCTGTGCTGCTGAGCCTGGAACTGCAGGAGCTGCCAGGTT
ATGGAAGGCGGCCAAGTGAATATAAGCCCCTTCGGCATTGGTACATGTGGTGTACCTT
AATTGGCCTCATCTGGACTAGTGAATTCATCCTGCGTGCCAGCAAATGACTATAAGCTGGGCAG
TGGTTACTTGTATTCAACAGAAGTAAAATGATCCTCCTGATCATCCCATTCTCGTCTCTC
TCCATTCTCTTCTACCATCAAGGAACCGTTGAAAGGGTATTAAATCTCTGTGGTGAG
GATTCCGAGAATCATGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCATTGT
CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTCTGGTGTCTGACAAATACCTGCTCCATCTC
AACCGAGATGCATATACTACAACGTCTATTAAATGGACAGATTCTGTACATCAGCAAAGATGC
ATTCAAATCTGTCCAAGAACTCAAGTCACTTACATCTATTAACTGCTTGGAGACTTCATAA
TTTTCTAGGAAAGGTGTTAGTGGTGTTCACTGTTGGAGACTCATGGCTTTACTAC
AATCGGGCATTCCAGGTGTGGCAGTCCCTCTGTTATTGGTAGCTTTGCCTACTTAGTGC
CCATAGTTTATCTGTGTTGAAACTGTGCTGGATGCACCTTCTGTGTTGCTGTTGATC
TGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTATGGATCAAGAATTCTGAGTTGTA
AAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGAATGAGGA
GGGAACAGAACTCCAGGCCATTGTGAGATAGATACCATTAGGTATCTGTACCTGGAAAACATT
TCCTCTAAGAGCCATTACAGAATAGAAGATGAGACCAGAGAAAAGTTAGTGAATT
TTAAAAGACCTAATAAACCTATTCTCCTCAAAA

FIGURE 32

MSGRDTILGLCILALALSLAMMFTFRFITTLLVHIFISIVILGLLFVCGVLWWLYYDYTNDLSIEDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLLFQPLWTFAIILFFWVLWAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQQMTIAGAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSLISVVRIPRIIVMYMQNALKEQQHGANLSRYLFRCYYCCFWCLDKYLLHLNQNAYTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFDGFIIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAHSFLSVFETVLDALFLCFAVDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIVR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 33.

GTTCGATTAGCTCCTCTGAGAAGAAGAGAAAAGGTTCTGGACCTCTCCCTGTTCTCCTTCTCTAGA
ATAATTGTATGGATTGTGATGCAGGAAAGCCTAAGGGAAAAAGAATATTCAATTCTGTGTGGT
GAAAATTTTGAACAAAGGAAAGGCTTCTCAAACAAGGGTGTCAATTCTGATATTATGAGGAC
TGTTGTTCTCACTATGAAGGCATCTGTTATTGAAATGTTCTTGTGCTGGTGACTGGAGTAC
ATTCAAACAAAGAACGGCAAAGAAGATTAAAAGGCCAAGTCACTGTGCCTCAGATCAACTGC
GATGTCAAAGCCGAAAGATCATGATCCTGAGTCATTGTGAAATGTCCAGCAGGATGCCAAGA
CCCCAAATACCATGTTATGGCACTGACGTGTATGCATCCTACTCCAGTGTGTGGCCTGCCG
TACACAGTGGTGTGCTGATAATTCAAGGAGGGAAAATACTGTTCGAAGGTTGCTGGACAGTCT
GGTTACAAAGGGAGTTATTCCAACGGTGTCCAATGTTATCCCTACCACGATGGAGAGAATCCTT
TATCGTCTTAGAAAGTAAACCCAAAAGGGTGTAAACCTACCCATCAGCTTACATACATCAT
CGAAAAGTCCAGCTGCCAAGCAGGTGAGACCACAAAAGCCTATCAGAGGCCACCTATTCCAGGG
ACAACACTGCACAGCCGGTCACTCTGATGCAGCTTCTGGCTGCACTGTAGCTGTGGCCACCCCCAC
CACCTGCCAAGGCCATCCCTCTGCTGCTTCTACCACCAGCATCCCCAGACCACAAATCAGTGG
GCCACAGGAGCCAGGAGATGGATCTCTGGTCCACTGCCACCTACACAAGCAGCCAAAACAGGCC
AGAGCTGATCCAGGTATCCAAGGCAAGATCCTCAGGAGCTGCCCTCCAGAAACCTGTTGGAGC
GGATGTCAGCCTGGACTGTTCCAAAAGAAGAATTGAGCACACAGCTTGGAGCCAGTATCCC
TGGGAGATCCAAACTGCAAATTGACTTGTGTTTAATTGATGGGAGCACCAGCATGGCAAA
CGGCGATTCCGAATCCAGAAGCAGCTCCTGGCTGATGTTGCCAAGCTTGTGACATTGCCCTGC
CGGTCCACTGATGGGTGTTGTCCAGTATGGAGACAACCCCTGCTACTCACTTTAACCTCAAGACAC
ACACGAATTCTCGAGATCTGAAGACAGCCATAGAGAAAATTACTCAGAGAGGAGGACTTCTAAT
GTAGGTCGGGCCATCTCTTGTGACCAAGAACTCTTCCAAAGCCAATGGAACACAGAAGC
GGCTCCCAATGTGGTGGTGTGATGGTGGATGGCTGGCCACGGACAAAGTGGAGGAGGCTTCAA
GACTTGCAGAGAGTCAGGAATCAACATTCTCATCACCATTGAAGGTGCTGCTGAAAATGAG
AAGCAGTATGTGGTGGAGCCAACTTGCAAACAAGGCCGTGCAAGAACAAACGGCTCTACTC
GCTCACGTGCAGAGCTGGTTGGCCTCCACAAGACCCCTGCAGCCTCTGGTGAAGCGGGCTGCG
ACACTGACGCCCTGGCCTGCAGCAAGACCTGCTGAACCTGGCTGACATTGGCTCGTCATCGAC
GGCTCCAGCAGTGTGGGACGGCAACTTCCGCACCGTCCAGTTGTGACCAACCTCACCAA
AGAGTTGAGATTCCGACACGGACACGCGCATCGGGGCCGTGCAGTACACCTACGAACAGCGGC
TGGAGTTGGGTCGACAAGTACAGCAGCAAGCCTGACATTCTCAACGCCATCAAGAGGGTGGGC
TACTGGAGTGGTGGCACCAGCACGGGGCTGCCATCAACTCGCCCTGGAGCAGCTCTCAAGAA
GTCCAAGCCAACAAGAGGAAGTTAATGATCCTCATCACCAGCAGGAGGTCTACGACGACGTCC
GGATCCCAGCCATGGCTGCCCATCTGAAGGGAGTGTACCTATGCGATAGGCCTGCTGGCT
GCCCAAGAGGAGCTAGAAGTCATTGCCACTCACCCGCCAGAGACCACCTCTTGTGGACGA
GTTTGACAACCTCCATCAGTATGTCCCCAGGATCATCCAGAACATTGTACAGAGTTCAACTCAC
AGCCTCGGAACTGAATTCAAGAGCAGGCAGAGCACCAGCAAGTGCTGTTACTAACTGACGTGTT
GGACCACCCACCGCTTAATGGGGCACGCACGGTGCATCAAGTCTGGGCAGGGCATGGAGAAC
AAATGTCTGTTATTATTCTTGCATCATGCTTTCATATTCCAAAATGTTAGTACAAAGA
TGATCACAAACGTATAGAATGAGCAGGCTACATCATGTTGAGGGTGTGGAGATTACAT
TTTGACAATTGTTCAAATGTTAGTGTGTTATTCTGATTTGAACCTCTGTAACCCCTCAGCAAGTT
AGCTTTGTGAGATTTAAGTGTGTTATTCTGATTTGAACCTCTGTAACCCCTCAGCAAGTT
TTTGTGATGACAATGTAGGAATTGCTGAATTGTTAGAAGGATGAAAATAAAAAA
AAG

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG
CQDPKYHVGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSLSLPRWR
ESFIVLESKPKKGVTYPSALTYSSSKSPAAQAGETTKAYQRPIPCTTAQPVTLMQLLAVTVAVA
TPTTLPRPSPSAASTTSIPRPQSVGHRSQEMDLWSTATYTSSQNRRADPGIQRQDPSGAAFQKP
VGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFRIQKQLLADVAQALDI
GPAGPLMGVVQYGDNPATHFNLKHTNSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
RSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVVEPNFANKAVCRTNG
FYSLHVQSWFGLHKTILQPLVKRVCCTDRLACSKTCLNSADIGFVIDGSSVGTGNFRTVLQFVTN
LTKEFEISDTDTRIGAVQTYEQRLEFGFDKYSSKPDILNAIKRVCYWSGGTSTGAAINFALEQL
FKKSCKPNKRKLMILITDGRSYDDVRIIPAMAAHLKGVITYAIGVAWAQEELEVIATHPARDHSFF
VDEFNDNLHQYVPRIIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCCTTAGGAGGTGGCTCGCTTGGAAAAGCTATCAAGGAAGAAATTGC
CAAACCATGTCTTTCTGTTTCAGAGTAGTCACAACAGATCTGAGTGTAACTAAGCATGGAAT
ACAGAAAACAACAAAAACTTAAGCTTAATTCATCTGAAATTCCACAGTTCTAGCTCCCTGGACCC
GGTTGACCTGTTGGCTCTTCCCGCTGGCTGCTCTATCACGTGGTGCCTCCGACTACTCACCCCGAGTGTA
AAGAACCTTCGGCTCGCGTGCCTCTGAGCTGCTGTGG**GATGGCCTCGGCTCTGGACTGTCCTCCGAGTA**
GGATGTCACTGAGATCCCTCAAATGGAGCCTCGCTGCTGTCACTCCTGAGTTCTTGTGATGTGGTAC
CTCAGCCTTCCCCACTACAATGTGATAGAACCGGTGAACGGATGTACTCTATGAGTATGAGCCGATTAA
CAGACAAGACTTCACTCACACTCGAGAGCATTCAAACACTGCTCTCATCAAAATCCATTCTGGTCATT
TGGTGCACCTCCCACCCCTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTACTTGGGGTGAACAAAGTCT
TGGTGGGGATATGAGGTTCTTACATTTCTTATTAGGCCAAGAGGCTGAAAGAACAAATGTTGGC
ATTGTCCTTAGAGGATGAAACACCTCTTATGGTGCATAATCCGACAAGATTTTAGACACATATAATA
ACCTGACCTTGAAAACCATTATGGCATTCAAGGTGGTAACGTGAGTTGCCCAATGCCAAGTACGTAATG
AAGACAGACACTGATGTTCAATACTGGCAATTAGTGAAGTATCTTAAACCTAAACCACTCAGA
GAAGTTTACAGGTATCCTCTAATTGATAATTATTCTATAGAGGATTTACCAAAACCCATATT
CTTACCAAGGAGTATCCTCAAGGTGTTCCCTCCACTGCAGTGGTTGGGTATATAATGCCAGAGAT
TTGGTGCCAAGGATCTATGAAATGATGGTCACGTAAAACCCATCAAGTTGAAGATGTTATGTCGGGAT
CTGTTGAATTATTAAAAGTGAACATTCAATTCCAGAACACAAATCTTCTTCTATAGAATCC
ATTGGATGTCTGTCAACTGAGACGTGTGATTGCAGCCATGGCTTCTCAAGGAGATCATCACTTT
TGGCAGGTATGCTAAGGAACACCACATGCCATT**TAAC**TTCACATTCTACAAAAGCCTAGAAGGACAG
GATACCTGTGGAAAGTGTAAATAAAAGTAGGTACTGTGGAAAATTCAATGGGAGGTCAGTGTGCTGGCT
ACACTGAACTGAAACTCATGAAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTCAGG
CCCTCAAAGATGATATGTGGAGGAATTAAATATAAGGAATTGGAGGTTTGCTAAAGAAATTAAATAGG
ACCAACAAATTGGACATGTCATTCTGTAGACTAGAATTCTAAAGGGTGTACTGAGTTAAAGCTCA
CTAGGCTGTAAAACAAACATGTAGAGTTTATTGAAACAATGTCAGTGAAGGTTGTGTA
TATCTTATGTGGATTACCAATTAAATATGTAGTTCTGTGTCAAAAACTCTTCAGTGAAGTTTGTGTA
CTGAACAAAATTTCACCTGTTGGTCATTAAAGTACTTCAGTGAAGATGTTGCAGTATTCACAGTTATT
ATTATTAAATTACTCAACTTGTGTTTAAATGTTGACGATTCAATACAAGATAAAAGGATAG
TGAATCATTCTTACATGCAAACATTCCAGTTACTTAAGTCAACTGATCAGTTATTGATAACATCACTCCA
TTAATGTAAAGTCATAGGTCAATTGCATATCAGTAATCTTGGACTTGTAAATATTACTGTGGT
AATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERNWMYFYEYEPIYRQDFHF
TLREHSNC SHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFFLLGQEAEKEDKMLA
LSLEDEHLLYGDIIHQDFLDTYNNLT LKTI MAFRWVTEFCPNAKYVMKTDTDVFINTGNLVKYLL
NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPKVFPYCSGLGYIMSRDLVPRIYEMMGHV
KPIKFEDVYVGICLNLLKVNIHIPEDTNLFFLYRIHLDVCQLRRVIAAHGFSSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

CGCTCGGGCACCAAGCCGGCAAGGAGGAGCTGGTTGCTGGACGCAGTTGGGGCTCACTTTCTCAGCTCCTCTCATC
TCGTCCTGCCAAGAGAGTACACAGTCATTAATGAAGCCTGCCCTGGAGCAGAGTGGAAATATCATGTGTCGGGAGTGCTGTG
AATATGATCAGATTGAGTGCCTGCCCCGGAAAGAGGGAAAGTCGTGGTTATACCATCCCTTGCTGCAGGAATGAGGAGAA
TGAGTGTGACTCCTGCCCTGATCCACCCAGGTTGACCATCTTGAAAATGCAAGAGCTGCCAAATGGCTCATGGGGGGT
ACCTTGGATGACTTCTATGTGAAGGGGTCTACTGTGCAGAGTGCCGAGCAGGCTGGTACGGAGGAGACTGCATGCATGTG
GCCAGGTTCTGCAGGCCAAAGGGTCAGATTTGTTGAAAGCTATCCCTAAATGCTCACTGTGAATGGACCATTATGC
TAAACCTGGGTTGTCACTCCAACTAAGATTGTCATGTTGAGTCTGGAGTTGACTACATGTGCCAGTATGACTATGTTGAG
GTTCGTGTGGAGACAACCGCAGATGGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGGCCAGCTCTATCCAGAGCATAG
GATCCTCACTCCACGTCCCTTCCACTCCGATGGCTCCAAGAATTGACGGTTCCATGCCATTATGAGGAGATCACAGC
ATGCTCCTCATCCCTGTTCCATGACGGCACGTGCGTCTTGACAAGGCTGGATCTTACAAGTGTGCCTGCTGGCAGGC
TATACTGGCAGCGCTGTGAAAATCTCCTGAAAGAAAGAAACTGCTCAGACCCCTGGGGCCAGTCATGGTACAGAAAA
TAACAGGGGGCCCTGGCTTATCAACGGACGCCATGCTAAAATTGGCACCGTGGTGTCTTCTTTGTAACAACCTCTATGT
TCTTAGTGGCAATGAGAAAAGAACTGCCAGCAGAATGGAGAGTGGTCAGGGAAACAGCCATCTGCATAAAAGCCTGCCGA
GAACCAAAGATTCAAGACCTGGTGAGAAGGAGGTTCTCCGATGCAGGTTCAAGTCAGGGAGACACCATTACACCAGCTAT
ACTCAGCGCCTTCAGCAAGCAGAAACTGCAGAGTGCCCTACCAAGAAGCCAGCCCTCCCTTGGAGATCTGCCATGGG
ATACCAAACATCTGCATACCCAGCTCCAGTATGAGTGCATCTCACCCCTCACCGCCCTGGCAGCAGCAGGAGGACATGT
CTGAGGACTGGGAAGTGGAGTGGCGGGCACCACCTGCATCCCTATCTGCGGGAAAATTGAGAACATCACTGCTCAAAGA
CCCAAGGGTTGCGCTGGCGTGGCAGGCATCTACAGGAGGACCAGCGGGTGCATGACGGCAGCCTACACAAGGGAGC
GTGGTTCTAGTCTGCAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGGCTGCCACTGTGTTACTGACCTGGGAAG
GTCACCATGATCAAGACAGCAGACCTGAAAGTTGGGAAATTCTACCGGGATGATGACCGGGATGAGAACACCAC
AGAGCCTACAGATTCTGCATTCATTCTGCATCCAACTATGACCCATCCTGCTTGATGCTGACATGCCATCCTGAAGCT
CCTAGACAAGGCCGTATCAGCACCCGAGTCCAGCCATCTGCCCTGCCAGTCGGGATCTCAGCACTCCCTCAGGAG
TCCCACATCACTGTGGCTGGCTGGAATGCTGGCAGACGTGAGGAGCCTGGCTCAAGAACGACACACTGCGCTCTGGG
TGGTCAGTGTGGTGGACTCGCTGCTGTGAGGAGCAGCATGAGGACCATGGCATCCAGTGAGTGTCACTGATAACATGTT
CTGTGCCAGCTGGAAACCCACTGCCCTCTGATATCTGCACTGCAGAGACAGGAGGCATGCCCTGTGCTCTCCGGGA
CGAGCATCTCTGAGGCCACGCTGGCATCTGATGGACTGGCAGCTGGAGCTATGATAAAACATGCAGCCACAGGCTCTCCA
CTGCCCTCACCAAGGTGCTGCTTTAAAGACTGGATTGAAAGAAATATGAAACCATGCTCATGCACCTTGAGAAG
TGTTTCTGTATATCCGTCTGTACGTGTGCAATTGCGTGAAGCAGTGTGGCCTGAAGTGTGATTGGCTGTGAACCTGGCT
GTGCCAGGGCTCTGACTTCAGGGACAAAACCTCAGTGAAGGGTAGACCTCCATTGCTGGTAGGCTGATGCCGCTCCA
CTACTAGGACAGCCAATTGAAAGATGCCAGGGCTGCAAGAAGTAAGTTCTCAAAGAACCCATATACAAACCTCTCCA
CTCCACTGACCTGGTGGCTTCCCAACTTCAGTTATCGAATGCCATCAGCTTGACCCAGGGAGATCTGGCTCATGAG
GCCCTTTGAGGCTCTCAAGTTCTAGAGAGCTGCCGTGGGACAGGCCAGGGCAGCAGAGCTGGGATGTGGTGCATGCC
TGTGTACATGCCACAGTACAGTCTGGCCTTTCCCATCTCTGTACACATTAAATAAGGGTTGGCTCT
GAACTACAA
AA

FIGURE 38

MELGCWTQLGLTFLQLLLSSLPREYT VINEACPGAEWNIMCRECCEYDQIECVCPGKREVVGYT
IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGTLDDFYVKGFYCAECRAGWYGGDCMRCGQ
VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDGQII
KRVCGNERPAPIQSIGSSLHVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC
ACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNNSYVLSGNE
KRTCQQNGEWSGKQPICIKACREP KISDLVRRVLPMQSRETPLHQLYSAAFSKQKLQSAPTK
KPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRTCLRTGKWSGRAPSCIICGKIENITAP
KTQGLRWPWQAAIYRRRTSGVHD GSLHKGAWFLVCS GALVNERTVVVAHCVTDLGKVTMIKTADL
KVVLGKFYRDDRDEKTIQSLQISAIILHPNYDPILL DADIAILKLLDKARI STRVQPI CLAASR
DLSTS FQESHITVAGWNVLADVRSPGFKN DTLRSGVVSVVDSLLCEEQHEDHGIPVSVTDNMFCA
SWEPTAPS DICTAETGGIAAVSFPG RASPEPRWLMGLVWSWSYD KTC SHRLSTAFTKVL PFKDWI
ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTCCTACATCCTCTCATCTGAGAACAGAGAGATAATCTCTTACGGGCCGTGATTATTAACGTGGCTTAATC
TGAAGGTTCTCAGTCAAATTCTTGTGATCTACTGATTGTGGGGCATGGCAAGGTTGCTAAAGGAGCTGGCTGG
TTTGGGCCCTTAGCTGACAGAAGGTGGCAGGGAGAATGCAGCACACTGCTGGAGAATGAAGGCCTCTGTTGC
TGGTCTTGCCTTGGCTCAGCCTGCTAACTACATTGACAATGTGGCAACCTGCACTTCTGTATTAGAACTCTGTA
AAGGTGCCTCCCACACGGCCTGACCAAAGATAGGAAGAGGCGCTACAAGATGGCTGTCCAGACGGCTGTGCGAGCC
TCACAGCCACGGCTCCCTCCCCAGAGGTTCTGCAGCTGCCACCACCTCCTTAATGACAGACGAGCCTGGCTAGACA
ACCCCTGCCTACGTGTCTCGCAGAGGACGGCAGCCAGCAATCAGCCAGTGGACTCTGGCCGGAGCAACCGAACTA
GGGCACGGCCCTTGAGAGATCCACTATTAGAACGAGATCATTAAAAAAATAATCGAGCTTGAGTGTCTCGAA
GGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCGACCAGGGCAGGGAAAATTCTGAAAACACCACTGCCCTGAAG
TCTTCCAAGGTTGATTCCAGATGGTGAATTACCAAGATCAAGATCAATCGAGTAGATCCCAGTGAAA
GCCTCTTATTAGGCTGGTGGAGGTAGCGAAACCCACTGGTCCATATCATTATCCAACACATTATCGTGTGGGG
TGATGCCAGAGACGGCCGGCTACTGCCAGGAGACATCATTCTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTC
ACAACATACGCTGTGCGTCTCCTGCCAGCCCTGCCAGGGCTGTGGCTGACTGTGATGCGTGAACAGAAGTCCGCA
GCAGGAACAATGGACAGGCCGGATGCCCTACAGACCCGAGATGACAGCTTCATGTGATTCTCAACAAAAGTAGCC
CCGAGGAGCAGCTTGAATAAAACTGGTGCAGCAAGGTGGATGAGCCTGGGTTTCACTTCAATGTGCTGGATGGCG
GTGTGGCATATCGACATGGCAGCTTGGAGAATGACCGTGTGTTAGCCATCAATGGACATGATCTCGATATGGCA
GCCCAAGAAAGTGCAGCTCATCTGATTCAAGGCCAGTGAAGACGTGTTCACCTCGTGTCCGCCAGGTTGGCAGC
GGAGCCCTGACATCTTCAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCCAGGGCAGGGAGAGGAGCAACA
CTCCCAAGCCCTCCATCCTACAATTACTGTGATGAGAACGGTAAATATCCTAAAGACCCGGTGAATCTCTCG
GCATGACCGTGCAGGGGAGCATCACATAGAACATGGGATTGCCTATCTATGTCATCAGTGTGAGCCGGAGGAG
TCATAAGCAGAGATGGAAGAATAAAACAGGTGACATTGTTGAATGTGGATGGGTCGAACGTGACAGAGGTGAGCC
GGAGTGAGGCAGTGGCATTATTGAAAAGAACATCATCCTCGATAGTACTCAAAGCTTGGAAAGTCAGAGTATGAGC
CCCAGGAAGACTGCAGCACCCAGCAGCCCTGGACTCCAACCACACATGGCCCAACCCAGTGAACGGTCCCCATCCT
GGGTGATGTGGCTGGAATTACACGGTGTGTATAACTGTAAAGATATTGTATTACGAAGAACACAGCTGGAAAGTC
TGGGCTCTGCATTGTAGGAGTTATGAAGAACATGGAAACAAACCTTTTCATCAAATCCATTGTTGAAGGAA
CACACGATACATGATGGAAGAACATGGGATTAGATGTGGTGAATTCTTCTTGCTGTCATGGTAGAAGTACATCAGGAATGA
TACATGCTTGGCAAGACTGCTGAAAGAACATTAAAGGAAGAACATTACTCTAAACTATTGTTCTGGCTGGCACTT
TTTATAGAATCAATGATGGTCAGAGGAAACAGAAAATCACAAATAGGCTAAGAAGTGAACACTATATTATC
TTGTCAGTTTATTTAAAGAACATGTCAGGAAAGTATGTCAGGAAAGTATGTCATCTAAATGAAAGCCAGTT
ACACCTCAGAAATATGATTCAAAAAAAATTAAACTACTAGTTTTTCAGTGTGGAGGATTCTCATTACTCTAC
AACATTGTTATTTCTATTCAATAAAAGCCCTAAACAACTAAATGATTGATTGTATAACCCACTGAATT
CAAGCTGATTAAATTAAATTGGTATATGCTGAGTCTGCCAAGGGTACATTATGCCATTAAATTACAGCT
AAAATTTTAAATGCATTGCTGAGAACGTTGCTTCAACAAAGAACATTTTCAAGAAGTAAATTTCAGAAGTAAA

FIGURE 40

MKALLLLVLPWLSPTYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPDGASLTATAPS
PEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDGRSNRTRARPFERSTIRSRSFKKINR
ALSVLRRTKSGSAVANHADQGRENSENTTAPEVFPRLYHLIPDGEITSIKINRVDPSESLSIRLV
GGSETPLVHIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLLRQPCQVLWLTVM
REQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGVFI FNVLDDGVAYRHG
QLEENDRVLAINGHDLRYGSPESAAHLIQASERRVHLVSRQRSPDIFQEAGWNSNGSWSPG
PGERSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHREWDLPIYVISVEPGGVISRDRGR
IKTGDILLNVDGVELTEVRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAPP
SDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG
RIRCGDILLAVNGRSTSGMIHACLRLLKELKGRITLTIVSWPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
467-473, 603-609

FIGURE 41

ACCAAGGCATTGTATCTTCAGTTGTCATCAAGTTGCAATCAGATTGGAAAAGCTCAACTTGAAGCTTT
CTTGCCCTGCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAACATGGGCTTCAACCTGACT
TTCCACCTTCTACAAATTCCGATTACTGTTGCTGTTGACTTGTGCCTGACAGTGGTTGGTGGC
CACCAAGTAACACTTCGTGGTGCCATTCAAGAGATTCTAAAGCAAAGGAGTTCATGGCTAATTCC
ATAAGACCCCTCATTTGGGAAGGGAAAAACTCTGACTAATGAAGCATTCCACGAAGAAGGTAGAACTT
GACAACGTCCCTCTGTCCTCCTACCTCAGAGGCCAGAGCAAGCTCATTTCAAACCAAGATCTCAC
TTTGGAAAGAGGTACAGGCAGAAAATCCAAAGTGTCCAGAGGCCGGTATGCCCTCAGGAATGTAAAG
CTTACAGAGGGTCGCCATCCTCGTCCCCACCGAACAGAGAACACCTGATGTACCTGCTGGAA
CATCTGCATCCCTCCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCACTCACCAGGCTGAAGG
TAAAAAGTTAATCGAGCCAAACTCTGAATGTGGCTATCTAGAACCCCTCAAGGAAGAAAATTGGG
ACTGCTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTACAAGTGTGAGGAG
CATCCCAAGCATCTGGTGGTGGCAGGAACAGCACTGGTACAGTTACAGTGGATATTTGG
GGGTGTTACTGCCCTAACGAGAGCAGTTTCAAGGTGAATGGATTCTCTAACAAACTACTGGGAT
GGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAACATGAAATTCCCGCCCTG
CCTGAAGTGGTAAATATAACATGGCTTCCACACTAGAGAACAGGAATGAGGTGAACGCAGAACG
GATGAAGCTTACACCAAGTGTACGAGTCTGGAGAACAGATGGTGAGTAGTTGTTCTTATAAAT
TAGTATCTGTGGAACACAATCCTTATATCAACATCACAGTGGATTCTGGTTGGTGCATGACCC
TGGATCTTTGGTGTGTTGGAAAGAACTGATTCTTGTGCAATAATTGGCCTAGAGACTTCAA
ATAGTAGCACACATTAAGAACCTGTTACAGCTCATTGTTGAGCTGAATTTCCTTTGTATTTCT
TAGCAGAGCTCCTGGTGTAGAGTATAAACAGTTGTAACAAGAACAGCTTCTTAGTCATTTGAT
CATGAGGGTTAAATATTGTAATATGGATACTTGAAGGACTTATATAAAAGGATGACTCAAAGGATAA
AATGAACGCTATTGAGGACTCTGGTGAAGGAGATTATTAAATTGAAGTAATATATTATGGGAT
AAAAGGCCACAGGAATAAGACTGCTGAATGTCTGAGAGAACCAAGAGTTGTTCTCGTCCAAGGTAGAA
AGGTACGAAGATAACTGTTATTCAATTCTGTACAATCATCTGTGAAGTGGTGGTGTCAAGGT
GAGAAGGCGTCCACAAAAGAGGGAGAAAAGGCACGAATCAGGACACAGTGAACCTGGGAATGAAGA
GGTAGCAGGAGGGTGGAGTGTGGCTGCAAAGGCAGCAGTAGCTGAGCTGGTGCAGGTGCTGATAGC
CTTCAGGGGAGGACCTGCCAGGTATGCCCTCAGTGTGATGCCACCAGAGAACATCTCTATTAGT
TTTAAAGAGTTTGTAAAATGATTGTACAAGTAGGATATGAATTAGCAGTTACAAGTTACAT
ATTAACATAATAAAATATGTCTATCAAACCTCTGTAGTAAATGTGAAAAGCAAAA

FIGURE 42

MGFNLTFHLSYKFRLLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLLIGKGKTLTN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKKFNRAKLLNVGYLEALKEENWDCFI FHDV
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGVTALSREQFFKVNGFSNNYWGWGGED
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGGACTGAGCTCCCAGATCTGGG
CCGCTTGCCTCCTGCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTCCCACAACAG
ACGGGACAACCTGCAGAGCTGCAACCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGATGCCCAT
GTTCCAGAGGCAGAGGAGGCAGACACCCACTTCCCCATCTGCATTTCTGCTGCGGCTGCTGTC
ATCGATCAAAGTGTGGATGTGCTGCAAGACGTAGAACCTACCTGCCCTGCCCGTCCCCCTCCC
TTCCTTATTATTCCCTGCTGCCAGAACATAGGTCTTCCAATAAAATGGCTGGTTCTTGT
TCCAAA
AAAAAAAAAAAAAAAAAAAAAAA

FIGURE 44

MALSSQIWAACLLLLLLASLTGSVFPQQTGQLAELQPQDRAGARASWMMPMFQRRRRDTHFPI
CIFCCGCCHRSKCGMCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTCCCCAACATGCCTCACCC
TCATCTATATCCTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGT
TCCGGTGGTGGGCCGTGACTTCCCCCTGAAGTCAAAGTAAAGCAAGTTGACTCTATTGTCTG
GACCTTCAACACAACCCCTTTGTCACCACAGCCAGAAGGGGGCACTATCATAGTGACCCAAA
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAATGAAAG
AAGAATGACTCAGGGATCTACTATGTGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCA
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGAAGAGGATGTGATT
TATACCTGGAAGGCCCTGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTC
CTGGAGATGGGAGAAAGTGTATGACCTTCATCTGCCTGCCAGGAACCCTGTCAGCAGAAACT
TCTCAAGCCCCATCCTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCCATG
GTCCTCCTGTCTCCTGTTGGTGCCCTCCTGCTCAGTCTCTTGTACTGGGCTATTCTTG
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTGTCGGG
AAACTCCTAACATATGCCCTATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
AGAACAACTCTAACAGGAAGATCCAGCAAATCGTTACTCCACTGTGGAAATACCGAAAAAGAT
GGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTGCCTATGAGAATGTTA
TCTAGACAGCAGTGCACTCCCCTAAGTCTGCTCA

FIGURE 46

MAGSPTCLTLIYILWQLTGSAAAGPVKELVGSVGGAVTFPLSKVKQVDSIVWTFNTTPLVTIQP
EGGTIIVTQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK
PKVTMGLQSNKNGTCVTNLTCMEHGEEDVIYTWKALGQAANESHNGSILPISWRWGESDMTFIC
VARNPVSRNFSSPILARKLCEGAADPDSSMVLLCLLVPLLLSLFVLGLFLWFLKRERQEYIE
EKKRVDICRETPNICPHSGENTEYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLLTMPDT
PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGGATGGACATCCTGCAATGG
ATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCCTCAATTGTCAGCT
TAGTTGAGGAAGACCAATTTCCTCAAAACCCCATCTCTGCTTGAGTGGTGGTCCCAGGAATT
ATAGGAGCAGGTCTGATGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAAGAGCGTG
CTGCAACAACAGAACTGGAATGTTCTTCATCATTTCAGTGTGATCACAGTCATTGGTGCTC
TGTATTGCATGCTGATATCCATCCAGGCTCTCTAAAAGGTCTCATGTGTAATTCTCCAAGC
AACAGTAATGCCAATTGTAATTTCATTGAAAAACATCAGTGACATTCCAGAATCCTTCAA
CTTGCAGTGGTTTCAATGACTCTTGTGCACCTCCTACTGGTTCAATAAACCCACCAGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTCCACTTCGATTCTGAAGAAAACAAACATAGG
CTTATCCACTCTCAGTATTAGGTCTATTGCTTGGAAATTCTGGAGGTCTGTTGGCT
CAGTCAGATAGTCATCGGTTCCCTGGCTGTGGAGTCTCAAGCGAAGAAGTCAAATTG
TGTAGTTAATGGGAATAAAATGTAAGTATCAGTAGTTGAAAAAA

FIGURE 48

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA
TTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNANCEFSL
KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRLIHFSVFLGL
LLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTGCCAAGGTGACCTCGCAGGACACTGGTGA
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
ATCCGTGGCTGCAGACCCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCTCGAACTGTGA
CATGGAGAGAGTGACCCCTGGCCCTCTCCTACTGGCAGGCCTGACTGCCTTGGAAAGCCAATGACC
CATTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
TGCGGAGGGCTCCTGCCATTGCTGGATCGCGCAGTTCTGAGTGGCAAATGCAAATACAAGAG
CAGCCAGAACGAGCACAGTCCTGTACCTGAGAAGGCCATCCCACTCATCACTCCAGGCTCTGCCA
CTACTTGCT**GAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCAGCACC**
TCCTCCCCCTGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCT
TTCTGATCAGGAGGCTTCTTATGAATTAAACTGCCACCACCCCTCA

FIGURE 50

MERVTLALLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLAIAGIAAVLSGKCKYKS
SQKQHSPVPEKAIPPLITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG
AGGACAGGGAGTCGAAGGAGGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGCAAGG
AGGAGACCCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGCTGGCAGAGATGAAGTTCCAG
GGGCCCCTGGCCTGCCTCCTGCTGGCCCTCTGCCTGGCAGTGGGAGGCTGGCCCTGCAGAG
CGGAGAGGAAAGCACTGGACAAATATTGGGAGGCCCTGGACATGCCCTGGAGACGCCCTGA
GCGAAGGGTGGAAAGGCCATTGGCAAAGAGGCCGGAGGGCAGCTGGCTCTAAAGTCAGTGAG
GCCCTGGCCAAGGGACCAGAGAAGCAGTTGGCACTGGAGTCAGGCAGGTTCCAGGCTTGGCGC
AGCAGATGCTTGGCAACAGGGCGGGAAAGCAGGCCATGCTCTGGAAACACTGGCACGAGA
TTGGCAGACAGGCAGAAAGATGTCATTGACACGGAGCAGATGCTGTCCGGCTCTGGCAGGG
GTGCCTGGCACAGTGGCTGGAAACTCTGGAGGCCATGGCATCTTGGCTCTCAAGGTGG
CCTTGGAGGCCAGGGCAATCCTGGAGGTCTGGGACTCCGTGGTCCACGGATAACCCG
GAAACTCAGCAGGCAGCTTGAATGAATCCTCAGGGAGCTCCCTGGGTCAAGGAGGCAATGGA
GGGCCACCAAACCTTGGACCAACACTCAGGGAGCTGTGGCCAGCCTGGCTATGGTCAGTGAG
AGCCAGCAACCAGAATGAAGGGTGCACGAATCCCCCACCCTGGCTCAGGTGGAGGCTCCAGCA
ACTCTGGGGAGGCAGCGGCTCACAGTCGGCAGCAGTGGCAGCAGCAGTGGCAGCAG
AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAACAGTGGCAGCAGCAGTGGCAGCAG
TGGCGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGCAGCAGCAGTGGCAGCAGTGGCAG
CCTCCTGGGATCCAGCACC GGCTCCTCCGGCAACCACGGTGGAGCAGGGAGGAAATGGA
CATAAACCCGGGTGTGAAAGCCAGGAATGAAGCCGGAGCAGGGAAATCTGGATTAGGG
CTTCAGAGGACAGGGAGTTCCAGCAACATGAGGGAAATAAGCAAAGAGGGCAATGCCCTTG
GAGGCTCTGGAGACAATTATGGGGCAAGGGTCAGCTGGGAGTGGAGGTGACGCTGTT
GGTGGAGTCAATACTGTGAACCTGAGACGTCTGGATGTTAACCTTGACACTTCTGGAA
GAATTAAATCCAAGCTGGTTCATCAACTGGATGCCATAAACAGGACCAGAGAAGCTCTC
GCATCCCGTGACCTCCAGACAAGGAGCCACCAGATTGGATGGAGCCCCACACTCCCTTAA
AACACCACCCCTCATCACTAATCTCAGCCCTGCCCTGAAATAACCTAGCTGCCCAACAAA
AA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 52

MKFQGPLACLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGAAGSKVS
EALGQGTREAVGTGVRQVPGFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGADAVRGSWQGVP
GHSGAWETSGGHGIFGSQGGLGGQGQGNGPGLGTPWVHGYPGNSAGSGMNPQGAPWGQGGNGGPPNF
GTNTQGAVAQPGYGSVRASNQNEGCTNPPSGSGGGSSNSGGSGSQSGSSGSNGDNNNGSSGGS
SSGSSSGSSSGSSGGSSGGSSGNSGGSRGDSGSESSWGSSTGSSSGNHGGSGGGNGHKGCEKPGNE
ARGSGESGIQGFRGQGVSSNMREISKEGNRLLGGSGDNYRGQGSSWGSGGDAVGGVNTVSETSPGM
FNFDTFWKNFKSKLGFINWDAINKDQRSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,
90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161,
159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224,
236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252,
253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285,
283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301,
298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329,
325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,
387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGGGACAAGCTGCTCCCACAGAAGGATGTCGCTGCTGAGCCTGCCCTGG
CTGGGCCTCAGACCGGTGGCAATGTCCCCATGGCTACTCCTGCTGCTGGTTGTGGCTCCTGGCT
ACTCGCCCCCATCCTGGCTTGACCTATGCCTCTATAACAACAGCCGCCGGCTCCAGTGTTC
CACAGCCCCAAAACGGAACGGTTTGACCTGGGCTGATCACTCCTACAGAGGAGGGC
TTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGGGCTTACGGTATGGCTGGTCCCAT
CATCCCCTCATCGTTATGCCACCCGTACACCATCCGGTCTATCACCAATGCCTCAGCTGCCA
TTGCACCCAAGGATAATCTCTCATCAGGTTCTGAAGCCCTGGCTGGAGAAGGGATACTGCTG
AGTGGCGGTGACAAGTGGAGGCCACCGTCGGATGCTGACGCCGCCCTCCATTCAACATCCT
GAAGTCCTATATAACGATCTCAACAAGAGTGCACAAACATCATGCTTGACAAGTGGCAGCACCTGG
CCTCAGAGGGCAGCAGTCGTCTGGACATGTTGAGCACATCAGCCTCATGACCTGGACAGTCTA
CAGAAATGCATCTCAGCTTGACAGCCATTGTCAGGAGAGGCCAGTGAATATATTGCCACCAT
CTTGGAGCTCAGTGCCTGTAGAGAAAAGAACGCCAGCATATCCTCCAGCACATGGACTTCTGT
ATTACCTCTCCATGACGGGCGGCCTCCACAGGGCCTGCCCTGGCATGACTTCACAGAC
GCTGTCATCCGGGAGCGCGTCGCACCCCTCCCCACTCAGGGTATTGATGATTTTCAAAGACAA
AGCCAAGTCCAAGACTTGGATTCATTGATGTGCTCTGCTGAGCAAGGATGAAGATGGGAAGG
CATTGTCAGATGAGGATAAGAGCAGAGGCTGACACCTCATGTTGGAGGCCATGACACCACG
GCCAGTGGCCTCTGGGTCTGTACAACCTGCGAGGCACCCAGAATACCAGGAGCGCTGCCG
ACAGGAGGTGCAAGAGCTTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGGCC
AGCTGCCCTTCCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCAGCTCCCTCATC
TCCCGATGCTGCACCCAGGACATTGTTCTCCCAGATGGCCAGTCATCCCCAAAGGCATTACCTG
CCTCATCGATATTATAGGGTCCATCACAAACCCAACTGTGTTGGCCGGATCCTGAGGTCTACGACC
CCTTCCGCTTGACCCAGAGAGAACAGCAAGGGAGGTACCTCTGGCTTTATTCCCTTCCGCA
GGGCCAGGAACACTGCATGGGCAGGCAGTCGCATGGCGAGATGAAAGTGGCCTGGCGTTGAT
GCTGCTGCACCTCCGGTTCTGCCAGACACACTGAGCCCCGAGGAAGCTGGAATTGATCATGC
GCGCCGAGGGCGGGCTTGGCTGCGGGTGGAGCCCTGAATGTAGGCTTGCAGTGACTTCTGAC
CCATCCACCTGTTTTGCAGATTGTCATGAATAACGGTGTGTC
AAA

FIGURE 54

MSLLSLPWGLGLRPVAMSPWLLLLVVGSWLLARILAWTYAFYNNCRLQCFQPPKRNWFGHLG
LITPTEEGLKDSTQMSATYSQGFTVWLGPPIIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP
WLGEGLLGGDKWSRHRRLTPAFHFNILKSYITIFNKSANIMLDKWQHLAGSSRLDMFEHI
SLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSDGRRFHRC
RLVHDFTDAVIRERRRLPTQGIDDDFKDKAKSKTLDFIGVLLSKDEDGKALSDEDIRAEADTF
MFGGHDTTASGLSWVLYNLARHPEYQERCRQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHNPTVWPDPFVYDPFRDPENSKGRSP
LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLEIMRAEGGLWLRVEPLN
VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTTGA
GCCTACTCGTTGATTGCAACTATCATGGTGCTGTTGTGTTGCACTTACCCGTGTTCTGCCT
TTGGTGGCATAACAAGGGACTTGCACTTATCTGCATTTGCAGTCTTGGCATTGACGTGG
TACAGCCTTCCTTCATACCATTGCAAGGGATGCTGTGAAGAAGTGTGTTGCCGTGTTGC
ATAATTCATGCCAGTTTATGAAGCTTGGAAAGGCACATGGACAGAAGCTGGTGGACAGTTT
GTAACATCTCGAAACCTCTGTCTTACAGACATGTGCCCTTATCTTGCAAGCAATGTGTTGCTT
GTGATTCGAACATTGAGGGTTACTTTGGAAGCAACAATACATTCTCGAACCTGAATGTCAGTA
GCACAGGATGAGAAGTGGGTTCTGTATCTTGAGTGGAAATCTTCCTCATGTACCTGTTCC
TCTGGATGTTGTCCCCTGAATTCCCATGAATACAAACCTATTCAACAGCAACAGCAAAAAAAA
AAA

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCASFVWHNKGALIFCILQSLALTWYSLSFIPFAR
DAVKKCFAVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGCACTGGAGCACCCAGCAGGCCAACATGCTCTGCTGCTG
TACGTGCCGGTCATCGGGGAAGCCCAGACCGAGTTCCAGTACTTGAGTCGAAGGGCTCCCTGCCAGCTGAAGTCC
ATTTCAAGCTCAGTGTCTCATCCCCCTCCAGGAATTCTCACCTACCGCAGTGGAAAGCAGAAAATTGTACAAGCT
GGAGATAAGGACCTTGATGGCAGCTAGACTTGAGAATTGTCCATTATCTCCAAGATCATGAGAAGAAGCTGAGG
CTGGTGTAAAGATTGGACAAAAGAATGATGGACGCATTGACGCCAGGAGATCATGCACTGCCCTGCCGGACTTG
GGAGTCAAGATATCTGAACAGCAGGAGAAAAAATTCTCAAGAGCATGGATAAAACGGCACGATGACCATCGACTGG
AACGAGTGGAGAGACTACCAACCTCCACCCGTGGAAAACATCCCCGAGATCATCCTCACTGGAAGCATTCCACG
ATCTTGATGTGGGTGAGAATCTAACGGTCCCGATGAGTCACAGTGGAGGAGGGCAGACGGGATGTGGTGGAGA
CACCTGGTGGCAGGGAGGTGGGGCAGGGGCCGTATCCAGAACCTGCACGGCCCCCTGGACAGGCTCAAGGTGCTCATG
CAGTCCATGCCCTCCCGCACAAACATGGCATTGTTGGCTTCACTCAGATGATTGGAGAAGGAGGGGCCAGG
TCACTCTGGCGGGCAATGGCATCAACGTCCTCAAATTGGCCCAATCAGCCATCAAATTGACATGGCTATGACAG
ATCAAGCGCTTGTGGTAGTGACCAAGGAGACTCTGAGGATTACGAGAGGCTTGTGGCAGGGCTTGGCAGGGCC
ATCGCCAGAGCAGCATCTACCAATGGAGGTCTGAAGACCCGGATGGCAGGCTGCGAAGACAGGCCAGTACTCAGGA
ATGCTGGACTGCGCCAGGAGGATCCTGGCCAGAGAGGGGGTGGCCGCTTCTACAAAGGCTATGCCCCAACATGCTG
GGCATCATCCCCATGCCGGCATGACCTTGCACTACGAGACGCTAAGAATGCCCTGGCTGCAGCACTATGAGTG
AACAGCGCGAACCCGGCTGTTGTGCTCCTGCCGTGGCACCAGTCCAGTACCTGTGCCAGCTGCCAGCTAC
CCCCCTGGCCCTAGTCAGGACCCGGATGCAGGCCAAGCCTCTATTGAGGGCGCTCCGGAGGTGACCATGAGCAGCCTC
TTCAAACATATCCTGCGGACCGAGGGGGCTTCGGGCTGTACAGGGGGCTGGCCCCAACTTCATGAAGGTATCCA
GCTGTGAGCATCAGCTACGTGGTCTACGAGAACCTGAAGATCACCTGGCGTGCAGTCGGTGAACGGGGAGGGC
CGCCCGCAGTGGACTCGCTGATCCTGGCCGAGCCTGGGTGTGCAGCCATCTCATTCTGTAATGTGCAACACT
AAGCTGTCTCGAGCCAAGCTGTGAAAACCTAGACGACCCCGCAGGGAGGTGGGAGAGCTGGCAGGCCAGGGCTT
GTCCTGCTGACCCAGCAGACCCCTCTGGTCCAGCGAAGACACAGGCATTCTAGGGTCCAGGGTCAGCAG
CTCCGGCTCACATGTGAAGGACAGACATTTCAGTCAGTGGCTGCAATAGTGAAGCTTGGAGCCTGGAGGCGGCT
TAGTTCTCCATTTCACCCCTGAGCCAGCTGTTGGCCACGGCCCTGCCCTCTGGTCTGCCGTGCATCTCCCTGTG
CCTCTGCTGCCCTGCTGAGGTAAAGTGGGAGGAGGCTACAGCCCACATCCCACCCCTGTCCAATCCC
ATAATCCATGATGAAAGGTGAGGTACGTGGCTCCCAGGCCTGACTTCCAACTACAGCATTGACGCCAATTGGC
TGTGAAGGAAGAGGAAAGGATCTGGCCTTGTGGTCACTGGCATCTGAGCCCTGCTGATGGCTGGGCTCTGGGATG
CTTGGGAGTGCAGGGGGCTGGGCTGCCCTGGCTGCAAGAGCAAGTGTGGGCTCATGGTCTGTGAGCT
GGCCTGGACCTGTCAGGATGGGCCCCACCTCAGAACCAACTCACTGCCCCACTGTGGCATGGGAGTGGAGCA
CCATTTGAGGGCGAAGGGCAGAGCCTTGTGTTCTGGGAGGGAAAGGAAAGGTGTTGGAGGCTTAATTATGG
ACTGTGGAAAAGGTTTGTCCAGAAGGACAAGCCGGACAAATGAGCAGCTCTGTGCTTCCAGAGGAAGACGAGG
GAGCAGGAGCTGGCTGACTGCTCAGAGTCTGACGCCCTGGGGTCTGTCAACCCAGCAGGGCGCAGC
GGGACCAAGCCCCACATTCACTTGTGCACTGCTGGACCTATTATTTGTATTATTAACAGAGTTATGTGCT
AACTATTTATAGATTGTTAATTAAAGCTGTCAAGTCATTTCATTTTATTCAATTATGTTCATGGTT
GATTGTACCTCCAAGCCGCCAGGGATGGGAGGAGGAGGAGAAGGGGGCTGGCCGCTGAGTCACATCT
GTCCAGAGAAATTCTTTGGACTGGAGGCAGAAAAGGCCAGAAGGCAGCAGGCCCTGGCTCCTTGGCAG
GTTGGGAAGGCTTGGCCCTAGGATTGACTGGGGCGTGGAGAGAGAGGGAGGAACCTCAAT
AACCTGAAAGGTGGAATCCAGTTATTCTGCGCTGCGAGGGTTCTTATTCACTCTTCTGAATGTCAAGGAG
TGAGGTGCTCTCACTGTGAATTGTGGTGGGAGGGGGCTGGAGGAGGAGGGTGGGGCTGGCTCCGTCCCTCCAGC
CTTCTGCTGCCCTGCTTAACAATGCCGCCAAGTGGCACCTCACGGTTGCACTCCATTCCACCAAGATGACCTGA
TGAGGAAATCTCAATAGGATGCAAAGATCAATGCAAAATTGTTATATGAACATATAACTGGAGTCGTCAAAAG
CAAATTAAAGAAATTGGACGTTAGAAGTTGTCAATTAAAGCAGCCTCTAATAAAAGTGTGTTCAAAGCTGAAAAAA
AA

FIGURE 58

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKDLG
QLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEKILKSMDKNG
TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMWWRHLVAGGG
AGAVSRTCTAPLDRLKVLMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESAIK
FMAYEQIKRLVGSQETLRIHERLVAGSLAGAIAQSSIYPMEVLKTRMALRKTGQYSGMLDCARR
ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNAWLQHYAVNSADPGVFVLLACGTMSSTC
GQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV
VYENLKITLEGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATA CGCTGGAACCTTCCCCAGCCATGGC
TTCCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
CACTCATCATGGCTTGGTATTCAGGGAGACACTCCATCACAGTCACTACTGTCGCCTCAGCT
GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTTGAACCTGACATCAAACATTCTGATAT
CGTGATACAATGGCTGAAGGAAGGTGTTAGGCTGGTCCATGAGTTCAAAGAAGGCAAAGATG
AGCTGTCGGAGCAGGATGAAATGTTAGGCTGGGACAGCAGTGGCTGATCAAGTGATAGTT
GGCAATGCCTCTTGCGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
CATCACTCTAAAGGCAAGGGGAATGCTAACCTGAGTATAAAACTGGAGCCTCAGCATGCCGG
AAAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTGCGGTGTGAGGCTCCCCGATGGTCCCC
CAGCCCACAGTGGTCTGGGATCCCAAGTTGACCAGGGAGCCAATTCTCGGAAGTCTCCAATAC
CAGCTTGAGCTGAACCTGAGAATGTGACCATGAAGGTTGTCTGTGCTCTACAATGTTACGA
TCAACAAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGATATCAAAGTG
ACAGAATCGGAGATCAAAAGGCGGAGTCACCTACAGCTGCTAAACTCAAAGGCTCTGTGT
CTCTCTTCTTGCCATCAGCTGGCACTTCTGCCTCTCAGCCCTACCTGATGCTAATAAT
GTGCCTTGGCACAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTCAC
CACCAGATATGACCTAGTTATATTCTGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTG
AGCAAACAAGAGCAAGAAACAAAAAGCCAAAGCAGAAGGCTCCAATATGAACAAAGATAAT
CTATCTCAAAGACATATTAGAAGTTGGAAAATAATTGATGTGAAGTGTGTTAAGA
GTGATAAGTAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGT
CACCTGGGAGTGAGAGGACAGGATAGTGCATGTTCTTGCTCTGAATTTAGTTATATGTGC
TGTAATGTTGCTCTGAGGAAGCCCTGGAAAGTCATCCAAACATATCCACATCTTATATTCCAC
AAATTAAAGCTGTAGTATGTACCTAACGCTGCTAATTGACTGCCACTCGCAACTCAGGGCG
GCTGCATTTAGTAATGGGTCAAATGATTCACTTTATGATGCTTCAAAGGTGCCTGGCTTC
TCTTCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTAGCATAAACAGAGCAGT
CGGGGACACCGATTTATAAAATAACTGAGCACCTCTTTAAACAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 60

MASLGQILFWSIISIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
DIVIQWLKEGVGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTDAGTYKC
YIITSKGKGNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFPQPTVVWASQVDQGANFSEVS
NTSFELNSENVTMKVSVLYNVTINNTYSCMIENDIAKATGDIKVTESI**KRRSHLQLLNSKASL**
CVSSFFAISWALLPLSPYLMK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTCAGAATCACCATGGCCAGCTATCCTACCGGCAGGGCTGCCAGGAGCTGCAGGACAAG
CACCAAGGAGCCCCTCCGGTAGCTACTACCCCTGGACCCCCAATAGTGGAGGGCAGTATGGTAGT
GGGCTACCCCTGGTGGTGGTATGGGGCTCTGCCCTGGAGGGCTTATGGACCACCAGCTGG
TGGAGGGCCCTATGGACACCCCAATCCTGGATGTTCCCTCTGGAACCTCAGGAGGACCATATG
GCGGTGCAGCTCCGGGGCCCTATGGTCAGCCACCTCCAAGTCCCTACGGTGCCAGCAGCCT
GGGCTTATGGACAGGGTGGCGCCCTCCAATGTGGATCCTGAGGCCTACTCCTGGTCCAGTC
GGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
ATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTGACAAGACCAAGTCA
GGCCGCATCGATGTCTACGGCTTCAGCCCTGTGGAAATTCCAGCAGTGGAAAGAACCTCTT
CCAGCAGTATGACCGGGACCGCTCGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGTCCC
AAATGGGCTACAACCTGAGCCCCAGTTCACCCAGCTCTGGTCTCCGCTACTGCCACGCTCT
GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGACAGA
GGCCTCCGGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTCGAGGACTCGTCA
CCATGACAGCTCTCGGATGCTTGACCAACCATCTGTGGAGAGTGGAGTGCACCAGGGACCTT
TCCTGGCTCTTAGAGTGAGAGAAGTATGTGGACATCTCTTTCTGTCCCTCTAGAAGAAC
ATTCTCCCTTGCTTGATGCAACACTGTTAAAAGAGGGTGGAGAGTCCTGCATCATGCCACCA
AATAGTGAGGACCGGGCTGAGGCCACACAGATAGGGCCTGATGGAGGAGAGGATAGAAGTTGA
ATGTCCTGATGGCCATGAGCAGTTGAGTGGCACAGCCTGGCACAGGAGCAGGTCTGTAATGG
AGTTAGTGTCCAGTCAGCTGAGCTCCACCCCTGATGCCAGTGGTGAGTGTTCATGGCCTGTTACC
GTTAGTACCTGTGTTCCCTCACCAAGGCATCCTGTCAAACGAGCCATTTCTCAAAGTGGAAAT
CTGACCAAGCATGAGAGAGATCTGTCTATGGGACCAAGTGGCTGGATTCTGCCACACCCATAAAAT
CCTTGTGTGTTAACTCTAGCTGCCTGGGCTGCCCTGCTCAGACAAATCTGCTCCCTGGCATT
CTTGCCAGGCTCTGCCCTGCAGCTGGACCCCTCACTGCCTGCCATGCTCTGCTCGGCT
TCAGTCTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTAATTGCATTTTTTC
ATTGGGCCAAAAGTCCAGTGAAATTGTAAGCTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYPGPPNSGGQYGSGLPPGGYGGPAPGGPYGPPAGGGPYG
HPNPGMFPSTPGGPYGGAAPGGPYGQPPPSSYGAQQPGLYQGGAPPNVDPEAYSWFQSVDSDH
SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTSGRIDVYGFSALWKFQQWKNLFQQYDR
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQVLTEAFREK
DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

FIGURE 64

MQGRVAGSCAPLGLLVCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGPSNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPPAGGSAVQRWPPSWGLPAMDSWPPEDPWQMMAAAEDRLGEA
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRLPRSNSLGAGGKILSQRP
PWSLIHRVLPDHPWGTLNPSVSWGGGGPGTGWGTRPMPHPEGIWGGINNQPPGTSGNINRYPGGS
WGNINRYPGGSWGNINRYPGGSWNIHLYPGINNPFPPGVLRPPGSSWNIPAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCAGGAGCGCAGTGGCCACTATGGGTC
TGGGCTGCCCTTGTCCCTCCTGACCCCTCCTGGCAGCTCACATGGAACAGGGCCGGGTATGA
CTTGCAACTGAAGCTGAAGGAGTCTTCTGACAAATTCCCTCATGAGTCCAGCTCCTGGAA
TTGCTTGAAAAGCTCTGCCTCCCTCCATCTCCCTCAGGGACCAGCGTCACCCTCACCATGC
AAGATCTAACACCATGTTGTCTGCAACACATGACAGCCATTGAAGCCTGTGTCTTGGCCC
GGGCTTTGGGCCGGGATGCAGGAGGCAGGCCCGACCCTGTCTTCAGCAGGCCACCCTC
CTGAGTGGCAATAAATAAAATTGGTATGCTG

FIGURE 66

MGSGPLVLLTLLGSSHGTGPGMTLQLKLKESFLTNSSYESSFELLEKLCLLLHLPSGTSVTL
HHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCGGGC
CAGGTGCCCGTCGCAAGTGCCCTGGCCGGAGATGCGTAGGAGGGCGAGCGCGAGAACCCCC
TTCCTCGCGCTGCCAACCCGCCACCCAGCCATGGCGAACCCCGGGCTGGGCTGCTTCTGGCG
CTGGGCCTGCCGTTCCCTGCTGGCCCGCTGGGGCCGAGCCTGGGGCAAATACAGACCACTCTGC
AAATGAGAATAGCACTGTTGCCTTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGTCTCTCCCTTGGCTGCCTGCTCCTGGCTGTGGGCTG
GCACTGTTGGTGGGAAGCTCGGGAGAACGGCAGACGGAGGGCACCTACCGGCCAGTAGCGA
GGAGCAGTTCTCCCATGCAGCCGAGGCCGGCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
GCCTGCCATCTAGGTCCCCTCTGCATCTGTCTCCCTCATTGCTGTGACCTTGGGAA
GGCAGTGCCCTCTGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAAGGTACTCAA
AGACTCTGCCCTGAGGTCAAGAGAGGATGGGGCTATTCACTTTATATATTATATAAAATTAG
TAGTGAGATGTAAAAAAAAAAAAAA

FIGURE 68

MANPGLGLLALGLPFLARWGRAWGQIQTTSANENSTVLPSSTSSSDGNLRPEAITAIIVVFS
LLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

GCAGGAATAACTAGAGAGGAACAATGGGTTATTCAGAGGTTTGTTCCTCTAGTCTGTGCCTGCTGCACCAAG
TCAAATACCTCCTCATTAAGCTGAATAATAATGGCTTGAAGATATTGTCAATTGTATAGATCCTAGTGTGCCAGAA
GATGAAAAAATAATTGAACAAATAGAGGATATGGTACTACAGCTTACGTACCTGTTGAAGCCACAGAAAAAGA
TTTTTTCAAAATGTATCTATATTAAATTCTGAGAATTGAAGGAAATCCTCAGTACAAAGGCCAAACATGAA
AACCATAAACATGCTGATTTAGTGCACCACTACACTCCCAGGTAGAGATGAACCATAACCCAAGCAGTTACA
GAATGTGGAGAGAAAGGCGAATACATTCACCTCACCTCCGGTGGGAGTGTGATGAGTACAATGAAGATCAGCCTTC
CCAGGCAAACTGTTGTCCATGAGTGGCTCACCTCCGGTGGGAGTGTGATGAGTACAATGAAGATCAGCCTTC
TACCGTCTAACTCAAAATCGAAGCAACAAGGTGTTCCGCAGGTATCTCTGGTAGAAATAGAGTTATAAGTGT
CAAGGAGGCAGCTGTTAGTAGAGCATGCAGAATTGATTCTACAACAAAACTGTATGGAAAAGATTGTCAATTCTT
CCTGATAAAGTACAAACAGAAAAGCATCCATAATGTTATGCAAAGTATTGATTCTGTTGAATTGTAAAGAA
AAAACCCATAATCAAGAAGCTCCAAGCCTACAAACATAAAAGTCAATTAGAGTACATGGGAGGTGATTAGCAAT
TCTGAGGATTTTAAACACCATACCCATGGTACACCCACTCCTCCACCTGTCTCTCATGCTGAAGATCAGTCAA
AGAATTGTGCTTAGTCTTGATAAGTCTGGAAGCATGGGGTAAGGACGCCCTAAATCGAATGAATCAAGCAGCA
AAACATTCTGCTGCAGACTGTTAAAATGGATCCTGGGGATGGTCACTTGATAGTACTGCCACTATTGTA
AATAAGCTAACAAATAAAAGCAGTGTGAAAGAAACACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGA
ACTTCCATCTGCTCTGGAATTAAATATGCATTCAGGTGATTGGAGAGCTACATCCCAACTCGATGGATCCGAAGTA
CTGCTGCTGACTGATGGGAGGATAACACTGCAAGTCTGTATTGATGAAGTGAACAAAGTGGGCCATTGTTCAT
TTTATTGCTTGGGAAGAGCGTGCTGATGAAGCAGTAATAGAGATGAGCAAGATAACAGGAGGAAGTCAATTGTT
TCAGATGAAGCTCAGAACAAATGGCCTCATTGATGCTTGGGCTCTACATCAGGAAUACTGATCTCCAGAAG
TCCCTCAGCTGAAAGTAAGGGATTAACACTGAATAGTAATGCCCTGGATGAACGACACTGTCAATTGATAGTACA
GTGGAAAGGACACGTTCTTCATCACATGGAACAGTCTGCCTCCAGTATTCTCTGGATCCAGTGGAAACA
ATAATGGAAAAATTTCACAGTGGATGCAACTTCCAAATGGCCTATCTCAGTATTCCAGGAAC
TGGCATACAACTTCAAGCCAAGCGAACCCAGAAACATTAAC
TGGCCTCCAATCACAGTGAATGCTAAATGAATAAGGACGTAAACAGT
ATTCTACAAAGGATATGTACCTGTTGGAGCCAATGTGACTGCTTCATTGAATCACAGAATGGACATACAGAAGTT
TTGGAACTTTGGATAATGGTGCAGGCGCTGATTCTTCAAGAATGATGGAGTCACTCCAGGTATTACAGCATAT
ACAGAAATGGCAGATATAGCTAAAGTCTGGGCTCATGGAGGAGCAAACACTGCCAGGCTAAATTACGGCCTCCA
CTGAATAGAGCCCGTACATACCAGCTGGTAGTGAACGGGAAATTGAAGCAAACCCGCCAAGACCTGAAATTGAT
GAGGACTCAGACCACCTGGAGGATTCCAGCGAACAGCATCCGGAGGTGCATTGTTGATCACAGTCCAAAGC
CTTCCCTGGCTGACCAATACCCACCAAGTC
ACATGGACAGCACCAGGAGATAATTGATGTTGGAAAGTTCAACGTTATATCATAAGAATAAGTGAAGTATTCT
GATCTAAGAGACAGTTGATGATGCTCTCAAGTAAATACTACTGATCTGTCACCAAGGAGGCCACTCCAAGGAA
AGCTTGCATTAAACAGAAATATCTCAGAAGAAATGCAACCCACATATTGTCATTAAAGTATAGATAAA
AGCAATTGACATCAAAGTATCCACATTGCAACAGTAACCTGTTATCCCTCAAGCAAATCCTGATGACATTGAT
CCTACACCTACTCCTACTCCTACTCCTGATAAAAGTCATAATTCTGGAGTTAATATTCTACGCTGGTATTG
TCTGTGATTGGGCTGTTGAATTGTTAACTTATTAAAGTACCCATTGAACTTAAGAAGAAAAAAATCTTC
AAGTAGACCTAGAAGAGGTTAAACAAACATGTAAGTAAAGGATATTGATTAATTTGATGAAATCTAAATTCA
GTGTGATCATAAAACTCATAAAATAATTAAAGATGTCGAAAAGGATACTTGTGATTAATTTGATGAAACAAAG
TGTAAAACGTCAAGATAAAATTAAATAGTTCAATTGTTATTGTTATTGTAAGAAATAGTGTGAAACAAAG
ATCCTTTTCAACTGATACTGGTGTATATTGATGCAACAGTTCTGAAATGATATTCAAATTGCATCAA
GAAATTAAACATCTATGAGTAGTC
AAAAAAGGAGGAGCAAATAACAAACATTGGAAAAAA
AAAAAAAAAA

FIGURE 70

MGLFRGFVFLVLCLLHQSNSTFIKLNNNGFEDIVIVIDPSVPEDEKIIIEQIEDMVTASTYLFE
ATEKRFKKNVSLIPENWKENPQYKRPKHENKHADIVVAPPTLPGRDEPYTKQFTECGEKG
IHFTPDLGGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRC
SAGISGRN
RVYKCQGGSCLSRA
CRIDSTTKLYGKDCQFPDKVQTEKASIMFMQSIDS
VVEFCNEKTHNQEAP
SLQNIKCNFRSTWEVISNS
EDFKNTIPMVTPPPPVFSLLKISQRIVCLVLDKSGSMGGKDR
LNR
MNQAAKHFLQTVENG
SWGMVHF
DSTATIVNKL
I
QIKSS
DERNT
L
MAGLPTYPLGGTS
ICSGIK
YAFQVIGELHSQ
LDGSEV
LLL
TDGEDNT
ASSCIDEV
KQSGAIVH
FIALGRA
ADEAVI
EMSKITGG
SHFYV
SDEAQNN
GLIDA
FGALTSG
NTDLSQ
KSLQ
LESKG
LTLNS
NAWM
NDT
VI
IDST
VGKDT
FFL
ITW
NSL
PPS
ISL
W
D
PSG
T
IMEN
FT
V
DAT
SK
MAY
L
S
I
PGT
AKV
GT
WAY
NLQ
AKAN
P
ETL
T
IT
V
TSR
AAN
SSV
P
P
ITV
NAK
MN
K
DV
NS
FP
SP
MIV
YAE
I
LQ
GY
V
P
V
L
G
A
N
V
TAF
IES
QNG
H
TEV
LE
L
D
N
G
A
G
D
SF
K
N
D
G
V
Y
S
R
Y
F
T
A
Y
T
E
N
G
R
Y
S
L
K
V
R
A
H
G
G
A
N
T
A
R
L
K
L
R
P
L
N
R
A
A
Y
I
P
G
W
V
V
N
G
E
I
E
A
N
P
P
R
E
I
D
E
D
T
Q
T
T
L
E
D
F
S
R
T
A
S
G
G
A
F
V
V
S
Q
V
P
S
L
P
L
D
Q
Y
P
P
S
Q
I
T
D
L
D
A
T
V
H
E
D
K
I
I
L
T
W
T
A
P
G
D
N
F
D
V
G
K
V
Q
R
Y
I
I
R
I
S
A
S
I
L
D
R
D
S
F
D
D
A
L
Q
V
N
T
T
D
L
S
P
K
E
A
N
K
S
E
F
A
K
P
E
N
I
S
E
E
N
A
T
H
I
F
I
A
I
K
S
I
D
K
S
N
L
T
S
K
V
S
N
I
A
Q
V
T
L
F
I
P
Q
A
N
P
D
D
I
D
P
T
P
T
P
T
P
D
K
S
H
N
S
G
V
N
I
S
T
L
V
L
S
V
G
V
I
I
S
T
T
I

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592, 628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCTTAGGTGAAACCTGGAGTAGAGTACTGACAGCAAAGACCGGAAAGACCATACTGCCCCGGGAGGGTGA
CAACAGGTGTATCTTTGATCTCGTGTGGCTGCCTCCTATTCAGGAAAGACCCAAGGTAATTTGACCCA
GAGGAGCAATGATGTAGCCACCTCTAACCTCCCTCTGAACCCCAAGTTATGCCAGGATTTACTAGAGAGTGTCA
ACTCAACCAGCAAGCGGCTCTCGGCTTAACCTGTGGTTGGAGGAGAGAACCTTGTGGGCTGCCTCTTAGCA
GTGCTCAGAAGTGAATTGCTGAGGGTGGACCAGAAGAAGGAAAGGCTCCCTCTGCTGTGGCTGCACATCAGGAA
GGCTGTGATGGGAATGAAGGTGAAAGCTGGAGATTCTACCTCAGTCATTGCTCTGCCTGCAAGATCATCCTTTAAA
AGTAGAGAAGCTGCTCTGTGGGGTTAACTCCAAGAGGAGAAGCTGTTCTAGAAGGAAATGGATGCAAGCAGCTC
CGGGGGCCCAAACGCATGTTCTGTGGCTAGCCCAGGGAAAGCCCTCCGTGGGGCCCGGCTTGAGGGATGCC
ACCGGTTCTGGACGCATGGCTGAATGATGATGATGGTTCGCCGGGGCTGCTGCGTGGGATTCCCGGGTGGT
GTTTGCTGGTGCCTCTGCTGTGCTATCTGCTGTACATGTTGGCCTGCACCCCAAAAGGTGACGAGGAGCAG
CTGGCACTGCCAGGGCAACAGCCCCACGGGAAGGAGGGTACAGGCCGCTCAGGAGTGGGAGGAGCAGCAC
CGCAACTACGTGAGCAGCCTGAAGCGCAGATGCCACAGCTCAAGGAGGAGCTGCAGGAGAGGAGTGCAGCTCAGG
AATGGCAGTACCAAGCCAGCGATGCTGCTGGCTGGACAGGGAGGGTACAGGCCACAGCTGAGGATATGCAGCAGTG
CTGGCCTTCCTGCACTCGCAGGTGGACAAGGCAGAGGTGAATGCTGGCGTCAAGCTGGCACAGAGTATGCAGCAGTG
CCTTCGATAGCTTACTCTACAGAAGGTGTACAGCTGGAGACTGGCCTTACCGCCACCCGAGGAGAAGCCTGTG
AGGAAGGACAAGCGGGATGAGTTGGTGAAGCATTGAATCAGCCTGGAGACCTGAACAACTCTGAGAGAACAGC
CCCAATACCGCTTACACGGCTCTGATTTCATAGAAGGGATCTACCGAACAGAAAGGGACAAGGGACATTGTAT
GAGCTCACCTCAAAGGGGACCACAAACACGAATTCAAACGGCTCATCTTATTCGACCATTAGCCCATCATGAAA
GTGAAAAATGAAAAGCTAACATGGCAACACGTTATCAATGTTATCGCTCTAGCAAAAAGGTGGACAAGTTC
CGGCAGTTATGAGGAGATGTGATTGAGCAGGATGGAGACTTCAGTCTCACTGTTGTTACTTTGGG
AAAGAAGAAATAATGAAGTCAAAGGAATACTTGAAAACACTTCAAAGCTGCCACTTCAGGAACCTTACCTTCATC
CAGCTGAATGGAGATTTCGAGGACTTGTGATGTTGGAGCCCCTGGAGGGAAAGCAACGTCCTTCTC
TTTTCTGTGATGTTGGACATCTACTTCACATCTGAATTCTCAATACGTGAGGCTGAATACACAGCCAGGGAGAAG
GTATTATCCAGTTCTTCAGTCAGTACAATCTGGCATAATATACGGCCACCATGATGCAGTCCCTCCCTGGAA
CAGCAGCTGGTCAAAGGAAGGAAACTGGATTGGAGAGACTTGGATTGGATGACGTGTCAGTATCGTCAGAC
TTCATCAATATAGGTGGTTGATCTGGACATCAAAGGCTGGGGCGGGAGAGGATGTGCACCTTATCGCAAGTATCTC
CACAGCAACCTCATAGTGGTACGGACGCCTGTGCGAGGACTTCCACCTCTGGCATGAGAACGCGCTGCATGGACAG
CTGACCCCCGAGCAGTACAAGATGTGATGCCAGGCAACGGGATCCCACGCCAGCTGGCATGCTG
GTGTTCAAGGAGATAGAGGCTCACCTCGAACAGAAACAGAACAGAACAGTAGCAAAAAACATGAACTCCCAGA
GAAGGATTGTGGGAGACACTTTCTTCCTTGCAATTACTGAAAGTGGCTGCAACAGAGAAAAGACTCCATAAA
GGACGACAAAAGAATTGGACTGATGGGTAGAGATGAGAAAGCCTCCGATTCTCTGTTGGCTTTACAACAGA
AATCAAAATCTCGCTTGCCTGCAAAGTAACCCAGTTGCACCCCTGTGAAGTGTCTGACAAAGGCAGAATGCTTG
AGATTATAAGCTTAAGAGCAGTTTGATGGTGGAGGTTTACAATACACTGAGAACCTGTTGCTGCTCATG
AATATTGATGTTAACAGGACTGTTCTAGGAATGCTAAATATCAGAACGGCAGGACATATTCTCCTCATATGAATGA
GCCTATCAGCAGGGCTCTAGTTCTAGGAATGCTAAATATCAGAACGGCAGGAGAGGAGATAGGCTTATTGATACT
AGTGAATGACATTAAGTAAAATGGACCAGAAAAGAAAAGAACATAAATATCGTGTATTTCCCAAGAT
TAACCAAAATAATCTGCTTATCTTGTGTTACTTAAATTACCACTTGCAGGCCCTACAAGAGAGCACAAGTGGCCTAC
TTTTTCCCTTGAGTTATGCTGCTTATTAATTACCACTTGCAGGCCCTACAAGAGAGCACAAGTGGCCTAC
ATTTTATTTTAAGAAGATACTTGAGATGCATTATGAGAACCTTCAGTCAGTCAGGACATGCTGAACTTGT
CCAAGGACATGCCAAATGCTGATTCTGTCAGGACTGAATGTCAGGACATAGGGAAAGGATGGTTGACT
AATACAGACGTACAGATACTTCTGAAAGAGTATTTCGAACAGGAGCAACTGAACACTGGAGGAAAAGAAAATGAC
ACTTCTGCTTACAGAAAAGGAAACTCATTCAAGACTGGTATCTGATGTCAGCTAAAGTCAAGAACACATTT
CTCCTCAGAAGTAGGGACCGCTTCTACCTGTTAAATAACCAAGTATACCGTGTGAACCAAACATCTCTTTC
AAAACAGGGTGCCTCTGGCTTCTGGCTTCCATAAGAAGAAATGGAGAAAATATATATATATATATATTGT
GAAAGATCAATCCATCTGCCAGAACACTAGTGGGATGGAAGTTTGCTACATGTTATCCACCCAGGCCAGGTGGAG
TAACGATTATTTTAAGCAGTTCTACTCAATCAGAACAGTCTGAAACATAGAGTGGTTCTCATTGATGAAATT
CAAACATTGTTAAAGGAGTAAACAGTTAACATAGAGTGGTTCTCATTGATGAAATTAGCCAGCACCAG
ATGCATGAGCTAATTATCTTGTGAGTCTGCTTGTGTTGCTCACAGTAAACTCATGTTAAAGCTTCAAGAAC
ATTCAAGCTGTGGTGTGTTAAAATGCATTGATTGATTGACTGGTAGTTATGAAATTAAATTAAACACAGG
CCATGAATGGAAGGTGGTATTGCACAGCTAATAAAATGATTGTGGATATGAA

FIGURE 72

MMMVRRGLLAWISRVVVLLVLLCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEWE
EQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVRKDKRDELVEAIESALETLNPA
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVKNELNMAN
TLINVIVPLAKRVDKFRQFMQNREMCEQDGRVHLTVVYFGKEEINEVKGILENTSKAANFRNF
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGKKVFYPVLFQSQY
NPGIIYGHDAVPPLSEQQLVIKKETGFWRDFGFGMTQCQYRSDFINIGGFDLDDIKGWGGEDVHLYR
KYLHSNLIVVRTPVRLFHLWHEKRCMDELTPEQYKMCMOSKAMNEASHGQLGMLVFRHEIEAHL
RKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 73

GAGACTGCAGAGGGAGATAAAGAGAGAGGGCAAAGAGGCAGCAAGAGATTGTCCTGGGATCCA
GAAACCCATGATAACCTACTGAACACCGAATCCCTGGAAGCCCACAGAGACAGAGACAGCAAGA
GAAGCAGAGATAAATACACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCTCACTCCTC
CCTCCCTCTCTCTGCCTGTCTAGTCCTCTAGTCCTCAAATTCCCAGTCCCCTGCACCCCTTC
CTGGGACACTATGTTCTCCGCCCTCTGCTGGAGGTGATTTGGATCCTGGCTGCAGATGGGG
GTCAACACTGGACGTATGAGGGCCCACATGGTCAGGACCATTGCCAGCCTCTTACCCCTGAGTGT
GGAAACAATGCCAGTCGCCATCGATATTCAAGACAGACAGTGTGACATTGACCTGATTGCC
TGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCGAGCCTTGGACCTGCACAACAATGCC
ACACAGTGCAACTCTCTGCCCTCACCCCTGTATCTGGGTGGACTTCCCCGAAAATATGTAGCT
GCCAGCTCCACCTGCACGGGTCAGAAAGGATCCCCAGGGGGTCAGAACACCAAGATCAACAG
TGAAGCCACATTGAGAGCTCCACATTGTACATTGACTCTGATTCTATGACAGCTTGAGTG
AGGCTGCTGAGAGGCCTCAGGGCCTGGCTGTCTGGCATCCTAATTGAGGTGGTGAGACTAAG
AATATAGCTTATGAACACATTCTGAGTCACCTGCATGAAGTCAGGCATAAAGATCAGAACACCTC
AGTGCCTCCCTAACCTAACAGAGAGCTGCTCCCCAACAGCTGGGCAGTACTTCCGCTACAATG
GCTCGCTACAACACTCCCCCTTGCTACCAGAGTGTGCTGGACAGTTTATAGAAGGTCCAG
ATTCAATGGAACACAGCTGGAAAAGCTTCAGGGACATTGTTCTCACAGAACAGGAGCCCTCTAA
GCTTCTGGTACAGAACTACCGAGCCCTCAGCCTCTCAATCAGCGATGGCTTGCTTCA
TCCAAGCAGGATCCTCGTATACCACAGGTGAAATGCTGAGTCTAGGTGAGGAATCTGGTGG
TGTCTGCCTCTCCTGGCTGTTATTCTAGAAAGATTGGAAGAACAGGCTGGAAAA
CCGAAAGAGTGTGGCTTCACCTCAGCACAAGCCACGACTGAGGCATAAAATTCTCTCAGATA
CATGGATGTGGATGACTCCCTCATGCCTATCAGGAAGCCTCTAAATGGGTGAGGATCTGG
CCAGAAACACTGTAGGAGTAGTAAGCAGATGTCCTCCTCCCTGGACATCTTAGAGAGGAAT
GGACCCAGGCTGTCATTCCAGGAAGAACTGCAGAGCCTCAGCCTCTCCAAACATGTAGGAGGAA
ATGAGGAAATCGCTGTGTTAATGCAGAGANAAACTCTGTTAGTTGCAGGGGAAGTTGGG
ATATAACCCAAAGTCCTCTACCCCTCACTTTATGCCCTTCCCTAGATATACTGCGGGATCT
CTCCTAGGATAAAGAGTTGCTGTTGAAGTTGATATTGATCAATATATTGAAATTAAAG
TTCTGACTTT

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWPASYPECGNNAQSPIDIQTDSVTFPDLPALQ
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQINSEAT
FAELHIVHYDSDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP
FNLRELLPKQLGQYFRYNGSLTPPCYQSVLWTVFYRRSQISMELQLEKLFSTEEEPSKLLV
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVLGVCLCLLLAVYFIARKIRKKRLENRKS
VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGCTGCCTGGCGGCCCTGGGACGGCAGTCCCTGTGTC
TCTGGTGGTTGCCTAACACCTGCAAACATCACCTCTTATCCATCAACATGAAGAATGCCTACA
ATGGACTCCACCAGAGGTCTCAAGGAGTTAAAGTTACTACACTGTGCAGTATTCAATCACA
ATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTCTGTTGCTGACAGCTCC
AGAGAAGTGAAGAGAAATCCAGAAGACCTCCTGTTCCATGCAACAAATATACTCCAATCTGA
AGTATAACGTGTCTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGACCAACCAC
ACGCTGGTGCTCACCTGGCTGGAGGCCAACACTCTTACTGCGTACACGTGGAGTCCTCGTCCC
AGGGCCCCCTGCCGTGCTCAGCCTCTGAGAACAGCAGTGTGCCAGGACTTGAAAGATCAATCAT
CAGAGTTCAAGGCTAAAATCATCTCTGGTATGTTGCCATATCTATTACCGTGTCTTT
TCTGTGATGGGCTATTCCATCTACCGATATCCACGTTGGCAAAGAGAACACCCAGCAAATT
GATTTGATTTATGAAATGAATTGACAAAAGATTCTTGTGCCTGCTGAAAAAAATCGTGT
ACTTTATCACCTCAATATCTGGATGATTCTAAAATTCTCATCAGGATATGAGTTACTGGGA
AAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCAGGAAACCTGAGGCCCTCAGGA
GGAAGAGGAGGTGAAACATTAGGGTATGCTCGCATTGATGAAATTGGACTCTGAAG
AAAACACGGAAGGTACTTCTCACCCAGCAAGAGTCCCTCAGCAGAACAAACCCCGGATAAA
ACAGTCATTGAATATGAATATGATGTCAGAACCACTGACATTGCGGGGCTGAAGAGCAGGA
GCTCAGTTGCAGGAGGAGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTGGCAG
TCTGGGCCGAAACGTTACAGTACTCATCACCCCTCAGCTCCAAGACTAGACCCCTGGCG
CAGGAGCACACAGACTCGGAGGAGGGCCGGAGGAAGAGCCATGACGACCCCTGGCGACTGGGA
TCCCCAAACTGGCAGGCTGTGTATTCCCTCGCTGTCCAGCTCGACCAGGATTAGAGGGCTGCG
AGCCTCTGAGGGGATGGCTCGGAGAGGAGGGTCTCTATCTAGACTCTAGGGAGCCGGCT
CCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTGAGGAATGGGGTTATA
TGTGCAGATGGAAAACTGCCAACACTTCCTTGCCTTGTGAAAGTTTCCAGTTGTCAGTGT
TCACCCCTTGATCCCAGCCATAAGTACCTGGATGAAAGAAGTTTCCAGTTGTCAGTGT
CTGTGAGAATTACTTATTCTCTATTCTCATAGCACGTGTGATTGGTCATGCATGTA
GGTCTCTAACATGATGGTGGCCTCTGGAGTCCAGGGCTGGCGTTGTTCTATGCAGAGAA
AGCAGTCATAAAATGTTGCCAGACTGGGTGCAGAATTATTCAAGGTGGGTGT

FIGURE 76

MSYNGLHQRFKELKLLTLCISISSQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIY
SNLKYNVSLNTKSNRTWSQCVTNHTLVLWLEPNTLYCVHVESFVPGPPRRAQPSEKQCARTLK
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILIYGNEFDKRFFVPAEK
IVINFITLNISDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEVVKHLGYASHLMEIFC
DSEENTEGTSLTQQESLSRTIPPDKTVIEYEYDVRTTDICAGPEEQELSLOEEVSTQGTLLESQA
ALAVLGPQTLQYSYTPQLQDLDPLAQEHDTSEEGPEEEPSTTLVDWDPQTGRLCIPSLSSFDQDS
EGCEPSEG DGLGEEGLLSRLYEEPAPDRPPGENETYLMQFMEEWGLYVQMen

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCACCGTCCCCAGGTCTGGCATCCTGCACTTGCTGCCCTCTGACAC
CTGGGAAGATGGCGGGCCGTGGACCTTCACCCTCTCTGTGGTTGCTGGCAGCCACCTTGATC
CAAGCCACCCCTCAGTCCCCTGCAGTTCTCATCCTCGGCCAAAAGTCATCAAAGAAAAGCTGAC
ACAGGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAGTGCATGC
GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGCAGCCCTGGTGAACACCGTCCTGAAGCACATC
ATCTGGCTGAAGGTCACTCACAGCTAACATCCTCCAGCTGCAGGTGAAGCCCTGGCCAATGACCA
GGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA
TCGTGGAGTTCCACATGACGACTGAGGCCAAGCCACCATCCGCATGGACACCCAGTGCAAGTGGC
CCCACCCGCCCTGGCCTCAGTGACTGTGCCACCAGCCATGGGAGCCTGCGCATCCAAGTGCTGTA
TAAGCTCTCCTCCTGGTGAACGCCCTAGCTAACAGGTCAACCTCTAGTGCCATCCCTGC
CCAATCTAGTGAAAACCAGCTGTGTCCCCTGATCGAGGCTTCCTCAATGGCATGTATGCAGAC
CTCCTGCAGCTGGTGAAGGTGCCATTCCCTCAGCATTGACCGTCTGGAGTTGACCTCTGTA
TCCTGCCATCAAGGGTGACACCATTCACTACCTGGGGCCAAGTTGTTGGACTCACAGGGAA
AGGTGACCAAGTGGTCAATAACTCTGCAGCTCCCTGACAATGCCACCCTGGACAACATCCCG
TTCAGCCTCATCGTGAGTCAGGACGTGGTGAAGCTGCAGTGGCTGCTGTGCTCTCCAGAAGA
ATTCAATGGCCTGTTGGACTCTGTGCTCCTGAGAGTGCCATGGCTGAAGTCAAGCATCGGGC
TGATCAATGAAAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGAC
ACTCCCGAGTTTTATAGACCAAGGCCATGCCAAGGTGGCCAAGTGATCGTGTGGAGTGTT
TCCCTCCAGTGAAGCCCTCCGCCCTTGTTGACCTGGCTGAAGCCAGCTCGGAAGCTCAGT
TTTACACCAAAGGTGACCAACTTAACTCAACTGAATAACATCAGCTCTGATCGGATCCAGCTG
ATGAACTCTGGATTGGCTGGTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCA
CTCCATCCTGCTGCCGAACCAGAAATGCCAAATTAGATCTGGGTCCCAGTGTCAATTGGTGAAGG
CCTTGGGATTGAGGCAGCTGAGTCCTCACTGACCAAGGATGCCCTGTGCTTACTCCAGCCTCC
TTGTGGAAACCCAGCTCCTGTCTCCAGTGAAGACTGGATGGCAGCCATCAGGGAAAGGCTGG
GTCCCTGAGCTGGAGTATGGGTGTGAGCTCTAGACCAATCCCTCTGCAATCAATAACACTTG
CCTGTGAAAAA

FIGURE 78

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELKDHNATSILQQLPLSAMREK
PAGGIPVLGSLVNTVLKHI IWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE
FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDTIQQLYLGAKLLDSQGKVT
KWFNNSAASLTMPPTLDNIPFSLIVSQDVVKAAVAAVLSPEEFMVLDSVLPEAHRLKSSIGLIN
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSEALRPLFTLGIEASSEAQFYT
KGDQLILNLNNISSLRDIQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKLRSGVPVSLVKALG
FEAAESSLTKDALVLTPASLWKPSSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCCATGGCCTCT
CTTGGCCTCCAACTTGTGGCTACATCCTAGGCCTCTGGGCTTTGGGCACACTGGTGCCAT
GCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCAGCATTGTGACAGCAGTTGGCTCT
CCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC
ACCCTCTGGCCTGCCGCTGACATCCAGGCTGCCAGGCCATGATGGTGACATCCAGTGAAT
CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGCATGAGATGCACAGTCTCTGCCAGGAATCCC
GAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTCATCCTGGAGGCCTCTGGGATT
ATTCCCTGTTGCCTGGAATCTCATGGATCCTACGGGACTTCTACTCACCACGGTGCCTGACAG
CATGAAATTGAGATTGGAGAGGCCTTTACTTGGCATTATTCTCCCTGTTCTCCCTGATAG
CTGGAATCATCCTCTGCTTCTGCTCATCCCAGAGAAATCGCTCCAACACTACGATGCCTAC
CAAGCCCCAACCTCTGCCACAAGGAGCTCCAAGGCCTGGTCAACCTCCAAAGTCAAGAGTGA
GTTCAATTCCCTACAGCCTGACAGGGTATGTGTGAAGAACCAAGGGCCAGAGCTGGGGGTGGCTG
GGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGCCACAGGTGAGGGACACTACCAACTGGATCGT
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTTGCCATTGGATTGAGCAAAGGCAGAAATGGGG
GCTAGTGTAAACAGCATGCAGGTTGAATTGCCAAGGATGCTGCCATGCCAGCCTTCTGTTTCC
TCACCTTGCTGCTCCCCGCCCTAAGTCCCCAACCCCTCAACTGAAACCCCATTCCCTTAAGCCA
GGACTCAGAGGATCCCTTGCCCTCTGGTTACCTGGGACTCCATCCCCAACCCACTAATCACA
TCCCACGTACTGACCCCTCTGTGATCAAAGACCCCTCTCTGGCTGAGGTTGGCTCTAGCTCATT
GCTGGGGATGGGAAGGAGAAGCAGTGGCTTGTCAGGATGCTCTAACCTACTTCTCAAGCTTC
CCTCCAAAGAAACTGATTGGCCCTGGAACCTCCATCCACTCTGTATGACTCCACAGTGTCCA
GACTAATTGTGCATGAACTGAAATAAAACCACCTACGGTATCCAGGAAACAGAAAGCAGGATG
CAGGATGGGAGGACAGGAAGGCAGCCTGGACATTAAAAAAATA

FIGURE 8o

MASLGLQLVGYILGLLGLLGTIVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLPADIQAAQAMMVTSAAISLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPVAWNLHGILRDFYSPLVPSMKEIGEALYLGISSLFSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSSPRPGQPPKVKEFNSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCACACCGTCCGCGCCTCTCCCTCTGCTGGACCTTCCTCGTCTCTCCATCTCTCCCTCCTTC
CCCGCGTTCTCTTCCACCTTCTCTTCCACCTAGACCTCCCTCCGCCCCCTGCCCCCTTCC
GCCACCCGCTGCTTCCTGGCCCTCTCCGACCCCGCTCTAGCAGCAGACCTCCCTGGGGTCTGTGG
GTTGATCTGTGGCCCTGTGCCTCCGTGTCCTTCGTCTCCCTCCGACTCCGCTCCCG
ACCAGCGGCTGACCTGGGAAAGGAGGTCCCGAGGTGAGGGCCTCTCCCTGCTGGGA
CTCGCCTGCTCTGGTCCCCCTGGACTCCCACGCTCGAGCCGCCAGACATGTTCTGCCTTT
CCATGGGAAGAGATACTCCCCCGCGAGAGCTGGCACCCCTACTTGGAGGCCACAAGGCCTGATGT
ACTGCCTGCCTGTACCTGCTCAGAGGGGCCCATGTGAGTTACCGCCTCCACTGTCCGCCT
GTCCACTGCCCGCAGCCTGTGACGGAGCCACAGCAATGCTGTCCCAAGTGTGTGGAACCTCACAC
TCCCTCTGGACTCCGGGCCCCACCAAAGTCCCTGCCAGCACAAACGGGACCATGTACCAACACGGAG
AGATCTTCAGTGCCCATGAGCTGTTCCCTCCGCCTGCCAACCAAGTGTGTGAGCTGCAGCTGC
ACAGAGGGCCAGATCTACTGCCGCCTCACAAACCTGCCAACCAAGGCTGCCAGCACCCCTCCC
ACTGCCAGACTCCTGCTGCCAAGCCTGCCAAAGATGAGGCAAGTGTGAGCAATGGATGAAGAGGACA
GTGTGCAGTCGCTCCATGGGTGAGACATCCTCAGGATCCATGTTCCAGTGTGCTGGAGAAAG
AGAGGCCCGGGCACCCAGCCCCACTGGCCTCAGCGCCCTCTGAGCTCATCCCTGCCACTT
CAGACCCAAAGGGAGCAGGCAGCACAACTGTCAAGATCGCTGAAGGAGAACATAAGAAAGCCT
GTGTGCATGGCGGGAGACGTACTCCCACGGGAGGTGTGGCACCCGGCTTCCGTGCCTCGGC
CCCTGCCCTGCATCCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCC
CACCGAGTACCCCTGCCGTACCCCGAGAAAGTGGCTGGAAAGTGTGCAAGATTGCCAGAGG
ACAAAGCAGACCCCTGCCACAGTGAGATCAGTTCTACCAGGTGTCCCAAGGCACCAGGGCGGGTC
CTCGTCCACACATCGGTATCCCAAGCCCAGACAACCTGCGTCGCTTGCCTGGAACACGAGGC
CTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAGATGAGGAAACTGAGGCTCAGAGAG
GTGAAGTACCTGGCCAAGGCCACACAGCCAGAATCTTCACTTGACTCAGATCAAGAAAGTCAG
GAAGCAAGACTTCCAGAAAGAGGCACAGCACTCCGACTGCTCGCTGGCCCCACGAAGGTCACT
GGAACGTCTCCTAGCCCAGACCCCTGGAGCTGAAGGTACGCCAGTCCAGACAAAGTACCAAG
ACATAACAAAGACCTAAAGTGCAGATATGAGCTGTATAATTGTTATTATATATTAATAAA
TAAGAAGTTGCATTACCCCTAAAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE
GAHVSCYRLHCPPVHCPQPVTPEPQQCCPKCVEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELP
PSRLPNQCVLCSCTEGQIYCGLTTCPPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGPGTAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKKACVHGGKTY
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTEYPCRHPEKVAGKCKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH
SQNLPLSDQESQEARNPERGTALPTARWPPRRSLERLPSPDPGAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTGTCTCGATGGAGTAGACTCTAGAACAGGCCAGTTGCCCTCCGCTACGCAGAGCCTCTCC
GTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCTCTCTAATCCATCCGTACCTCTCCTGTCA
TCCGTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCTCTCATGCTCAGTTGGTCTGAGTC
TCCTCAAGCTGGATCAGGGCAGTGGCAGGTGTTGGGCCAGACAAGCCTGTCCAGGCCTGGTGGGGAG
GACGCAGCATTCTCCTGTTCCCTGTCTCTAAGACCAATGCAGAGGCCATGGAAGTGCAGGTTCTCAGGGG
CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACCAGCCATTATGCAGATGCCACAGTATC
AAGGCAGGACAAAACTGGTGAAGGATTCTATTGCGGAGGGGCCATCTCTGAGGCTGGAAAACATTACT
GTGTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCCCAGTCTTACTACCAGAAGGCCATCTGGAGCT
ACAGGTGTCAGCACTGGCCTCAGTTCTCTCATTCCATCACGGATATGTTGATAGAGACATCCAGCTAC
TCTGTCAGTCCTCGGGCTGGTCCCCGGCCACAGCGAAGTGGAAAGGTCCACAAGGACAGGATTGTCC
ACAGACTCCAGGACAAACAGAGACATGCATGGCCTGTTGATGTGGAGATCTCTGACCGTCCAAGAGAA
CGCCGGGAGCATATCCTGTTCCATGCGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAG
GAGATACCTTTGAGCCTATATCGTGGCACCTGGTACCAAAGTACTGGAAACTCTGCTGTGGCTA
TTTTTGGCATTGTTGACTGAAGATTTCTCTCAAATTCCAGTGGAAAATCCAGGCCAACTGGACTG
GAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAG
AGACGGCTACCCGAAGCTCTGCGTTCTGATCTGAAACTGTAACCCATAGAAAAGCTCCCCAGGAGGTG
CCTCACTCTGAGAAGAGATTACAAGGAAGAGTGTGGTGGCTCTCAGAGTTCCAAGCAGGGAAACATTA
CTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGGTGGAGTGTGCCGGATGATGTGGACAGGAGGA
AGGAGTACGTGACTTTGTCTCCGATCATGGTACTGGTCCCTCAGACTGAATGGAGAACATTGTATTTC
ACATTAATCCCCGTTTATCAGCGTCTTCCCAGGACCCCACCTACAAAATAGGGTCTCCTGGACTA
TGAGTGTGGACCATCTCCTCTCAACATAATGACCAGTCCCTTATTATACCCTGACATGTCGGTTG
AAGGCTTATTGAGGCCCTACATTGAGTATCCGTCTATAATGAGCAAATGGAACCTCCATAGTCATCTGC
CCAGTCACCCAGGAATCAGAGAAAGAGGCCTTGGCAAAGGCCCTGCAATCCCAGAGACAAGAACAG
TGAGTCCTCCTCACAGGCAACCACGCCCTCCTCCCCAGGGTGAAATGTAGGATGAATCACATCCCACAT
TCTCTTGGATATTAAGGTCTCTCTCCAGATCAAAGTCCCGCAGCAGCCGCCAAGGTGGCTTCCA
GATGAAGGGGACTGCCGTCCACATGGGAGTCAGGTGTCAGGCTGCCCTGAGCTGGAGGGAAAGG
CTGACATTACATTAGTTGCTCTCACTCCATCTGGCTAAGTGACTTGAAATACCACCTCAGGTGAAG
AACCGTCAGGAATTCCCATCTCACAGGCTGTGGTAGATTAAGTAGACAAGGAATGTGAATAATGCTTAG
ATCTTATTGATGACAGAGTGTATCCTAATGGTTGTCATTATATTACACTTCAGTAAAAAA

FIGURE 84

MALMLSLVSLLLKGSGQWQVFGPDKPVQALVGEDAASFCLSPKTNAEAMEVRFGRQFSSVH
LYRDGKDQPFMOMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISSQSYYQKAIWELQ
VSALGSVPLISITGYVDRDIQLLCQSSGWFPRPTAKWKGPGQDLSSTDRTNRDMHGLFDVEISL
TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
FQWKIQAELDWRRKHGQAEELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAPQEVPHSEKRF
TRKSVVASQSFOAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLSPDHGYWVLRLNGEHLYFT
LNPRFISVFPPRTPKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGT
PIVICPVTQESEKEASWQRASAIPETSNSESSQATTPFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGCGGCCCTGGCACCTAAACCCAGACATGCTGCTGCTGCTGCCCT
GCTCTGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTTCCGTGA
CGGTGCAGGAAGGCCTGTGTCCATGTGCCCTGCTCCTCTCCTACCCCTCGCATGGCTGGATT
TACCCCTGGCCCAGTAGTCATGGCTACTGGTCCGGAAAGGGGCAATACAGACCAGGATGCTCC
AGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACCGATTCCACCTCCTG
GGGACCCACATACCAAGAATTGCACCCCTGAGCATCAGAGATGCCAGAAGAAGTGTGATGCCGGGAGA
TACTTCTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCACCGGCTCTGTGAA
TGTGACAGCCTTGACCCACAGGCCAACATCCTCATCCCAGGCACCCCTGGAGTCCGGCTGCC
AGAATCTGACCTGCTCTGTGCCCTGGCCTGTGAGCAGGGACACCCCTATGATCTCCTGGATA
GGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCGCTCCTCGGTGCTCACCTCATCCCACA
GCCCCAGGACCATGGCACCAGCCTCACCTGTCAAGGTGACCTCCCTGGGCCAGCGTGACCACGA
ACAAGACCGTCCATCTCAACGTGCTCACCGCCTCAGAACTTGACCATGACTGTCTCCAAGGA
GACGGCACAGTATCCACAGTCTGGAAATGGCTCATCTGTCACTCCCAGAGGGCAGTCT
GCGCCTGGTCTGTGCAAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA
GAGGCCTGACCCCTGTGCCCTCACAGCCCTCAAACCCGGGGTGCTGGAGCTGCCTGGTGCAC
CTGAGGGATGCAGCTGAATTCACCTGCAGAGCTCAGAACCCCTCGGCTCTCAGCAGGTCTACCT
GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGTGGTCGGGGAGCTGGAG
CCACAGCCCTGGTCTCCTGCGTCATCTCGTTAGTGAGGTGCAGGAAGAAA
TCGGCAAGGCCAGCAGCGGGCGTGGGAGATAAGGGCATAGAGGATGCAAACGCTGTCAGGGTTC
AGCCTCTCAGGGGCCCTGACTGAACCTTGGCAGAAGACAGTCCCCAGACCAGCCTCCCCAG
CTTCTGCCGCTCTCAGTGGGGAGAGAGCTCCAGTATGCATCCCTCAGCTCAGATGGTG
AAGCCTTGGACTCGCGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG
ATGAGAAACTGCAGAGACTCACCCCTGATTGAGGGATCACAGCCCTCCAGGCAAGGGAGAAGTCA
GAGGCTGATTCTGTAGAATTAAACAGCCCTAACGTGATGAGCTATGATAACACTATGAATTATG
TGCAGAGTGAAAGCACACAGGCTTAGAGTCAAAGTATCTAACCTGAATCCACACTGTGCC
TCCCTTTTTTTAACTAAAAGACAGACAAATTCTA

FIGURE 86

MLLLLPLLWGRERAEGQTSKLLTMQSSVTQEGLCVHVPSCFSYPSHGWIYPGPVHGYWFREG
ANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTL SIRDARRSDAGRYFFRMEKGSIKWNY
KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCSPWACEQGTPPMISWIGTSVSPLDPSTTRS
SVLTLIPQPQDHGTSLTCQVTFP GASVTTNKTVHLNVSYPPQNLTMVFQGDGTVSTVLNGSSL
SLPEGQSLRLVCAVDADVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDAAEFTCRAQNP
LGSQQVYLNVLQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKK SARPAAGVGDTGIE
DANAVRGSASQGPLTEPWAEDSPPDQPPPASARSSVGEGELOQYASLSFQMVKPWDSRGQEATDTE
YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGTGAAGGGAGCTCTGTAC
CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTCTGCTGTTCTC
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT
TCTTCGTCTCCATCTGCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCTAGTGCATTGA
TGGCCTGTATTTCTCCGACTGAGAATGGTGTATCTACCAGACCTCTGTGACATGACCTCTG
GGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGAAAGTGCACGGTG
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGACGGCAACTGGC
CAACTACAACACCTTGGATCTGCAGAGGCAGGCCACGAGCGATGACTACAAGAACCCCTGGCTACT
ACGACATCCAGGCCAAGGACCTGGCATCTGGCACGTGCCAATAAGTCCCCATGCAGCACTGG
AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTCCTCAGACACTGGACATAATCT
GTTTGGCATCTACCAGAAATATCCAGTGAATATGGAGAAGGAAAGTGTGGACTGACAACGGCC
CGGTGATCCCTGTGGCTATGATTTGGCGACGCCAGAAAACAGCATCTTATTACTCACCCTAT
GGCCAGCGGAATTCACTGCAGGGATTGTCAGTTCAAGGTATTAATAACGAGAGAGCAGCAA
CGCCTGTGTGCTGGAATGAGGGTCACCGATGTAACACTGAGCATCACTGCATTGGAGGAG
GATACTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTCTGGTTTGATTGGAGTGGATAT
GGAACTCATGTTGGTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG
TTGAGAGTTGTGGAGGGAACCCAGACCTCTCCTCCACCAGAGATCCAAAGGATGGAGAA
CAACTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAATCATATTGACTCAAGA
AAAAAA

FIGURE 88

MNQLSFLFLIAATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRSSQQGSKADYPEGDGNWANYNTFGSAEAAT
SDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLRYRTDTGFLQTLGHNLFGIYQKYPVKYG
EGKCWTDNGPVIKVYDFGDAQKTASYYSPYQREFTAGFVQFRVFNNERAANALCAGMRVTGCN
TEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTIVGYSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTGTCGGCTTGC~~GGGG~~GAGACTCAGGAGTCGCTGTCTGAACCTCCAGCCTCAGAGAC
CGCCGCCCTGTCCCCGAGGGCCATGGGCCGGTCTCAGGGCTTGTGCCCTCGCTCCTGACG
CTCCTGGCGCATCTGGTGGT~~CGT~~CATCACCTATTCTGGTCCC~~GGG~~ACAGCAACATA~~CAGGC~~CTG
CCTGCCTCTCACGTTCACCCCCGAGGGAGTATGACAAGCAGGACATT~~CAG~~CTGGTGGCC~~CG~~CTCT
CTGTCACCCTGGGCCTTTGCAGTGGAGCTGGCCGGTTCC~~T~~TCAGGAGTCTCCATGTTAAC
AGCACCCAGAGCCTCATCTCCATTGGG~~G~~CTCACTGTAGTGCATCCGTGGCC~~CT~~GTCC~~TT~~CCAGCTG
TCACTGAAATGGCTTATT~~CG~~TACCGTCTTGGG~~G~~CTGAAAAAGAAAC~~C~~CTTGATTAC~~TT~~CA
TGACGGAACCTAAGGACGAAGCCTACAGGGCAAGGGCCGCTCGTATT~~C~~CTGGAAAGAAGGAAG
GCATAGGCTCGGTTCCC~~C~~CGAAACTGCTCTGCTGGAGGATATGTGTTGGAATAATTACG
TCTTGAGTCTGGGATTATCCGCATTGTATTTAGTGCTTGTAA~~AAA~~ATGTTGTAGTAACA
TTAAGACTTATACAGTTAGGGACAATT~~AAAAAAA~~AAAAAA

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSALPAVTEMALFV
TVFGLKKPFF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

FIGURE 91

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATGCAACTGAC
TCGCTGCTGCTCGTGTTCCTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGATGATG
GTCCTCCGGCTCAGAGGACCTGAGCGTGTGACGACACGAGGGCCAGCCCCGGCCCCGGTGCCT
CGGAAGCGGGGCCACATCTCACCTAAGTCCCACCCATGCCAATTCCACTCTCCTAGGGCTGCT
GGCCCCGCCTGGGGAGGGCTTGGGCATTCTTGGCAGCCCCCAACCGCCGAACCACAGCCCC
CACCTCAGCCAAGGTGAAGAAAATCTTGGCTGGGCACCTACTCCAACATCAAGACGGTG
GCCCTGAACCTGCTCGTCACAGGGAAAGATTGTGGACCATGGCAATGGACCTTCAGCGTCCACTT
CCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCCTCGTGCCTCCAGTAAAGCTGTAG
AGTCCACCAGGAACAGCAGATCTCATCGAACGCCAAGGCCTCCAAATCTCAACTGCCGGATG
GAGTGGGAGAAGGTAGAACGGGCCGGACCTCGCTTGACCCACGACCCAGCCAAGATCTG
CTCCCGAGACCACGCTCAGAGCTCAGGCCACCTGGAGCTGCTCCCAGCCCTCAAAGTCGTCTGTG
TCTACATGCCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAAC
CATAGTGATAACCCCTACTACCCATCTGGGTGACCCGGGGCAGGCCACAGAGGCCAGGGC
TGGAAGGACAGGCCCTGCCCATGCAGGAGACCATCTGGACACCAGGGCAGGGAAAGGGTTGGCCTC
AGGCAGGGAGGGGGGTGGAGACGAGGAGATGCCAAGTGGGCCAGGGCAAGTCTCAAGTGGCAG
AGAAAGGGTCCAAGTGCTGGTCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGG
AGGAGGAGTGGCTCTGTGCAGCCTCACAGGGCTTGCCACGGCCACAGAGAGATGCTGG
TCCCCGAGGCCGTGGCAGGCCATCAGTGTGGCCCCAGATCAAGTCATGGAGGAAGCTAAGC
CCTGGTTCTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGAGATTCAAGTGTGGACA
GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCCTAGGAGCCAGTCAGCAGGGTGGGGTGG
GCCAGAGGAGCTCCAGCCCTGCCTAGTGGCGCCCTGAGCCCTGTGCTGTGAGCATGG
CATGAGGCTGAAGTGGCAACCCCTGGGTCTTGATGTCTTGACAGATTGACCATCTGTCTCCAGC
CAGGCCACCCCTTCCAAAATTCCCTTTCTGCCAGTACTCCCCCTGTACCAACCCATTGCTGATG
GCACACCCATCCTAAGCTAACAGACAGGACGATTGTGGCTCCACACTAACGCCACAGCCCAC
CGCGTGCTGTGTCCCTCTCCACCCCAACCCCTGCTGGCTCTGGAGCATCCATGTCCCG
GAGAGGGGTCCCTAACAGTCAGCCTCACCTGTCAGACCGGGTTCTCCGGATCTGGATGGCG
CGCCCTCTCAGCAGCGGGCACGGTGGGCGGGCGGGCGAGAGCATGTGCTGGATCTGTT
TGTGTGTCTGTGTGGGTGGGGAGGGAGGGAAAGTCTGTGAAACCGCTGATTGCTGACTTT
TGTGTGAAGAATCGTGTCTTGGAGCAGGAAATAAGCTGCCCGGGCA

FIGURE 92

MQLTRCCFVFLVQGS¹LYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISP²KSRPMANST³L
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHGN⁴TF
SVHFQHNATGQGNIS⁵ISLVPPSKA⁶VEFHQEQQI⁷FIEAKASKIFNCRM⁸EWEKVERGRR⁹TS¹⁰LCTHDP
AKICSRDHAQSSATWSCSQPFKVV¹¹CVYIAFYSTDYRLVQKVCPD¹²NYHSDTP¹³YPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCCATGACTGCGGCCGTGTTCTCGGCTGCGCCTTCATTGCCCTCGGGCTGCGCTCGCC
CTTATGTCTCACCATGCCATCGAGCCGTTGCGTATCATCTCCTCATGCCGGAGCTTCTT
CTGGTTGGTGTCTACTGATTCGTCCTGTTGGTTATGGCAAGAGTCATTATTGACAACA
AAGATGGACCAACACAGAAATATCTGCTGATCTTGGAGCGTTGTCTGTCTATATCCAAGAA
ATGTTCCGATTGCATATTATAAACTCTTAAAAAAAGCCAGTGAAGGTTGAAGAGTATAAACCC
AGGTGAGACAGCACCCCTATGCGACTGCTGGCCTATGTTCTGGCTGGGCTTGAATCATGA
GTGGAGTATTCCTTGTGAATACCTATCTGACTCCTGGGCCAGGCACAGTGGCATTCA
GGAGATTCTCCTCAATTCTCCTTATTCACTGACTTCATGACGCTGGTCATTATCTGCTGCATGT
ATTCTGGGCATTGTATTTGATGGCTGTGAGAAGAAAAAGTGGGCATCCTCCTTATCGTC
TCCTGACCCACCTGCTGGTGTCAAGCCCAGACCTCATAAGTTCTTATTATGGAATAAACCTGGCG
TCAGCATTATAATCCTGGTGCATGGCACCTGGCATTCTAGCTGCAGGAGCTGCCG
AAGCCTGAAACTCTGCCTGCTGCCAAGACAAGAACTTCTTACAACCAGCGCTCCAGAT
AACCTCAGGGAACCGACACTCCAAACCGCAGACTACATCTTAGAGGAAGCACAACGTGCCT
TTTCTGAAAATCCCTTTCTGGTGAATTGAGAAAGAAATAAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAFGPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDNKDG
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV
FSFVNTLSDSLGP GTVGIHG DSPQFFLYSAFMTLVIIILLHVFWGIVFFDGCEKKW GILLIVLLT
HLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSLKLCLLCQDKNFLYNQRSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTCACCAGAGTAAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTGCCTCGTG
GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTGTCTTAGGATCAAC
TCGGTCATTACCACAGCTAAACCTGCTTGGACTCCCTCCCACAAAACTGGCTCCGGATCAGG
GAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTCCTTTAAGTCTGATACCATTAAACA
CAGATGCTCACACTGGGCCAGATCTGCATCTGTTAAATCCTGCTGCAGGAATGACACCTGGTAC
CCAGACCCACCCATTGACCCCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC
CAATTGGTACACAACTGGAGCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC
TTCACGAGCCTCATCATCCATTCTGTTCCGGAGGCATCCTGCCACCAGTCAGGCAGGGC
TAATCCAGATGTCCAGGATGGAAGCCTCCAGCAGGAGGAGCAGGTGTAATCCTGCCACCCAGG
GAACCCCAGCAGGCCCTCCAACTCCCAGTGGCACAGATGACGACTTGCAGTGACCACCCCT
GCAGGCATCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAAATGGAATTCA
GTAAGCTGTTCAAATTGGAACTAAGCTGCCTCGAATTGGTGATACTGTGAATCTTATC
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAAATTCTTAATT
TACCTGAAAATATTCTGAAATTCAAGAAAATATGTTCTATGTAGAGAATCCAACTTTAAAAAA
CAATAATTCAATGGATAAAATCTGTCTTGAATATAACATTATGCTGCCTGGATGATGCATAT
TAAAACATATTGGAAAACTGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAA

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSSLIPLTQM
LTLPDLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTILSSEE
LPQIIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:

amino acids 1-16

FIGURE 97

GCTCAAGTGCCTGCCTGCCACCCAGCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGCTCT
CTTCTGCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGGCGCTGGAGGGCTGTCCTGACCATG
GTCCTGCCTGGCTGTGGCTGCTTGTCTCCGTCCCCCAGGCTCTCCCCAAGGCCAGCCTG
AGAGCTGTCTGTGGAGTTCCAGAAAATGGTGGAAATTCCCTTATACCTGACCAAGTTGC
CGCTGCCCGTGAAGGGGCTGAAGGCCAGATCGTGTCAAGGGACTCAGGAAGGAACTGAG
GGCCATTGCTATGGATCCAGATTCTGGCTTCTGCTGGTACCGAGGCCCTGGACCGAGAGGA
GCAGGCAGAGTACAGCTACAGGTACCCCTGGAGATGCAGGATGGACATGTCTGTGGGTCAC
AGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTGCCCCATTCTCAAGCCATCTAC
AGAGCTGGCTGAGCCGGGTACCGGCCATCCCCCTCCTCTTGGTACAGGTCAAGGACATGGGTGA
GGATGAGCCAGGCACAGCCAACCTGGATCTCGATTCCACATCTGAGCCAGGCTCCAGCCCAGC
CTTCCCCAGACATGTTCCAGCTGGAGCCTCGGCTGGGCTCTGGCCCTCAGCCCCAAGGGGAGC
ACCAGCCTGACCACGCCCTGGAGAGGACCTACAGCTGTGGTACAGGTCAAGGACATGGGTGA
CCAGGCCTCAGGCCACCAGGCCACTGCCACCGTGGAAAGTCTCCATAGAGAGCACCTGGGTGA
CCCTAGAGCCTATCCACCTGGCAGAGAATCTAAAGTCTTACCCGCACCACATGGCCAGGTA
CACTGGAGTGGGGTGTGACTATCACCTGGAGAGGCCATCCCCGGGACCCCTTGAAGTGAA
TGCAGAGGAAACCTTACGTGACCAGAGACTGGACAGAGAAGGCCAGGCTGAGTACCTGCTCC
AGGTGCGGGCTCAGAATTCCATGGCAGGGACTATGCGGCCCCCTGGAGCTGCACGTGCTGGTGA
ATGGATGAGAATGACAACGTGCTTACCTGGCAGGGACCTGGGAGGAGACAGTCAGCATCCCTGAGCT
CAGTCCACCAGGTACTGAAGTGAATGACTAGACTGTCAGCAGAGGATGCAGATGCCCGGCTCCCCCA
ATTCCCACGTTGTATCAGCTCTGAGCCCTGAGCCTGAGGGATGGGTAGAGGGAGAGCCTTC
CAGGTGGACCCACTTCAGGCAGTGTGACGCTGGGGCTCCACTCCGAGCAGGCCAGAACAT
CCTGCTCTGGTCTGCCATGGACCTGGCAGGCCAGAGGGTGGCTCAGCAGCACGTGTAAG
TCGAAGTGCAGTCACAGATATCAATGATCACGCCCTGAGTTCATCACTTCCAGATGGCCT
ATAAGCCTCCCTGAGGATGTGGAGCCGGACTCTGGTGGCATGCTAACAGCCATTGATGCTGA
CTCGAGCCGCCCTCCGCTCATGGATTGCCATTGAGAGGGAGACACAGAAGGGACTTTG
GCCTGGATTGGAGCCAGACTCTGGCATGTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAG
GCAGCTCCAAGTCATGAGGTGGTGGTGGCAGAGTGTGGGAAGCTGGGGCAGGCC
AGGCCCTGGAGCCACGCCACGGTGAECTGTGCTAGTGGAGAGAGTGTGCCACCCCCAAGTTGG
ACCAGGAGAGCTACGAGGCCAGTGTCCCCATCAGTGCCTGGCTCTTCTGCTGACCAC
CAGCCCTCCGACCCATCAGCCGAACCCCTCAGGTCTCCCTAGTCATGACTCAGAGGCTGGCT
CTGCATTGAGAAATTCTCCGGGAGGTGCACACGCCAGTCCCTGCAGGGGCCAGCCCTGGGG
ACACCTACACGGTGTGTGGAGGCCAGGATACAGCCCTGACTCTGCCCCCTGTGCCCTCCAA
TACCTCTGCACACCCGCCAAGACCATGGCTGATCGTGAGTGGACCCAGCAAGGACCCGATCT
GCCAGTGGCACGGTCCCTACAGCTCACCTGGTCCAACCCACGGTGCACGGGATTGGC
GCCTCCAGACTCTCAATGGTCCCATGCTTACCTCACCTGGCCCTGCATTGGTGGAGCCACGT
GAACACATAATCCCCGTGGTGGCAGCCACAATGCCAGATGTGGCAGCTCTGGTTCAGTGAT
CGTGTGCGCTGCAACGTGGAGGGCAGTGCATGCGCAAGGTGGGCCAGTGAAGGGCATGCCA
CGAAGCTGTGGCAGTGGCATCTGTAGGCACCTGGTAGCAATAGGAATCTCCTCATCTC
ATTTCACCCACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCAGCAGACAGCGTCCCC
GAAGGCGACTGTCTGAATGGCCAGGCAGCTAGCTGGAGCTTGGCCTCTGGCTCCATCTGAG
TCCCCCTGGAGAGAGGCCAGCACCAAGATCCAGCAGGGACAGGACAGAGTAGAAGGCCCTCCA
TCTGCCCTGGGGTGGAGGCACCATCACCCTCACCAAGGCATGTCTGCAGAGCCTGGACACCAACTT
TATGGACTGCCATGGAGTGCTCAAATGTCAGGGTGTGTTGCCAATAATAAGGCCAGAGAA
CTGGCTGGCCCTATGGAAAAAAAAAAAAAAAAAAAAAG

FIGURE 98

MVPAWLWLLCVSVPQALPKAQPAELSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSGKAT
EGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLHVKDENDQVPHFSQAI
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQPSPD MFQLEPRLGALALSPKG
STSLDHALERTYQLLVQVKDMGDQASGHQATATVEVSIIESTWVSLEPIHLAENLKVLYPHHMAQ
VHWSSGGDVHYHLESHPPGPFEVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDYAAPLELHVL
VMDENDNVPICPPRDPTVSIPELSPPGTEVTRLSAEDADAPGSPNSHVYQLLSPEPEDGVEGRA
FQVDPTSGSVTLGVPLRAGQNI LLVILAMDLAGAEGGFSTCEVEVAVTDINDHAPEFITSQIG
PISLPEDVEPGTLVAMLTAIDADLEPAFRIMDFAIERGDTEGTFG LDWE PDSGHVRLRLCKNLSY
EAAPSHEVVVVVQSVAKLVGPGPGATATVTVLVERVMPPPQLDQESYEASVPISAPAGSFLLT
IQPSDPISRTLRFSLVNDSEGWL CIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALT LAPVPS
QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGPNPTVQRDWRLQTLNGSHAYLT LALHWEP
REHIIIPVVVSHNAQM WQLLVRVIVCRCNVEGQCMRKVGRMKGMPTKLSAVGILVGTIVAI GIFI
LIFTHWTMSRKKDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCACGCCCTG
AGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCACGCTCTGGAAGCACCAGCCTTA
TCTCTTACCTTCAAGTCCCCTTCTCAAGAATCCTCTGTTCTTGCCTCTAAAGTCTGGTAC
ATCTAGGACCCAGGCATCTGCTTCCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAAATG
TTCTCCTATGTTGGTCACTATTGCATTAGAAGCTGCAACAAATTCCAATGAGACTAGCACC
TCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACGCCACCAACTCTGGTCCAG
TGTGACCTCCAGTGGGTCAAGCACAGCCACCATCTCAGGGTCCAGCGTACCTCCAATGGGTCA
GCATAGTCACCAACTCTGAGTCCATACAACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACCTCCAGTGG
GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCTCAGTGGGCCAGCACAGTCACCAACT
CTGGTCCAGTGTGACCTCCAGTGGAGCCAGCAGTGCACCAACTCTGAGTCCAGCACAGTGTCC
AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCCAGTGGGCCAGCACAGCCAC
CAACTCTGACTCCAGCACAACTCCAGTGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGCCACCAACTCTGAGTCCAG
GCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAG
AACGACCTCCAATGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGCCA
GCACAGCCACCAACTCTGACTCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG
TCCAGCACGACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG
GGCTAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGCCAGCACAGCCACCAACT
CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACACCCTCC
AGTGGGCCAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGCCAACACAGCCAC
CAACTCTGAGTCCAGCACAGTGTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGTCAAGCACAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGCTAGCACA
GCCACCAACTCTGACTCCAGCACAACTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCTAG
CACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACAAACCTCCAGTGGGCCA
ACACAGCCACCAACTCTGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTGACTGGA
ATGCACACAACCTCCATAGTCATCTACTGCAGTGAGTGAGGCAAAGCCTGGTGGTCCCTGGT
GCCGTGGAAATCTCCTCATCACCCTGGTCTCGGTTGTGGCGGGCTGGGCTTTGCTGGC
TCTCTTCTGTGAGAAACAGCCTGTCCTGAGAAACACCTTAACACAGCTGTCTACCACCT
CATGGCCTCAACCATGGCTTGGTCCAGGCCCTGGAGGGAATCATGGAGCCCCCACAGGCCAG
GTGGAGTCCTAACTGGTCTGGAGGAGACAGTATCGATGCCATGGAGATGAGCAGGGAGGA
ACAGCAGGCCCTGTAGCAGCCCCGAAGCAAGTGCCGATCTCAGGAAGGAAGAGACCTGGCA
CCCAAGACCTGGTTCTTCATTCATCCCAGGAGACCCCTCCAGCTTGTGAGATCCTGAA
AATCTGAAGAAGGTATTCTCACCTTCTTGCTTACAGACACTGGAAAGAGAATACTAT
TGCTCATTAGCTAAGAAATAACATCTCATCTAACACACAGACAAAGAGAAGCTGTGCTTG
CCCCGGGGTGGGTATCTAGCTCTGAGATGAACACTGAGTTATAGGAGAAAACCTCCATGCTGGACTC
CATCTGGCATTCAAATCTCACAGTAAAATCAAAGACCTCAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MKMKGKVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATISGS
SVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTSSGASTATNSESSTPSS
GASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTSSGASTA
TNSESSTSSGASTATNSESSTVSSRASTATNSESSTSSGASTATNSESRTTSNGAGTATNSES
STTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTSSGASTATNSDSSTTSS
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA
TNSESSTTSSGVSTATNSESSTSSGASTATNSDSSTSSEASTATNSESSTVSSGISTVTNSES
STTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGSLVPWEIFLITLVVA
AVGLFAGLFFCVRNSLISLRNTFNTAVYHPHGLNHGLPGPGGNHGAPHPRWSPNWFWRPVSSI
AMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCGGGACGCCTCCCGCTTACGGGATGAATTAACGGCGGGTCCGCACGGAGGTTGTGACCCCTA
CGGAGCCCCAGCTGCCACGCACCCCACTCGCGTCGCGCGCGTGCCTGCTGTACAGGTG
GGAGGCTGGAACTATCAGGCTGAAAAACAGAGTGGGTACTCTCTGGGAAGCTGGCAACAAAT
GGATGATGTGATATATGCATTCCAGGGAAAGGGAAATTGTGGTGCTCTGAACCCATGGTCAATT
AACGAGGCAGTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCTTGGAAATCAT
GGTGTCAAGGAAAGGGATTACTTACTGACTCTGTTGGGAAGCTTTGGAAAGCATT
TCATGCTGAGTCCCTTTACCTTGATGTTGAAACCCATCTGGTATCGCTGGATCAACAAAC
CGCCTTGTGGCAACATGGCTACCCCTACCTGTGGCATTATTGGAGACCATGTTGGTAAAAGT
GATTATAACTGGGATGCATTGTTCTGGAGAAAGAAGTGTCAATTATCATGAACCATCGGACAA
GAATGGACTGGATGTTCTGTGGAAATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAAATT
TGCCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTGGTGGCCATGCAGGCTGCTGCCTATAT
CTTCATTCAAGGAAATGGAAGGATGACAAGAGGCCATTCGAAGACATGATTGATTACTTTGTG
ATATTCAACGAACCACTCAACTCCTCATATTCCAGAAGGGACTGATCTCACAGAAAACAGCAAG
TCTCGAAGTAATGCATTGCTAAAAAAATGGACTTCAGAAATATGAATATGTTACATCCAAG
AACTACAGGCTTACTTGTGGTAGACCGTCTAAGAGAAGGTAAGAACCTTGATGCTGTCCATG
ATATCACTGTGGCGTATCCTACAACATTCTCAATCAGAGAACCTCCCTCCAAGGAGACTTT
CCCAGGGAAATCCACTTCACGTCCACCGGTATCCAATAGACACCCTCCCCACATCCAAGGAGGA
CCTTCAACTCTGGTGCCACAAACGGTGGAGAGAGAAGAGAGAGGCTGCGTTCTTCTATCAAG
GGGAGAAGAATTTTACCGGACAGAGTGTCAATTGCAAGTCTGAACCTAGGGTC
CTTGTGGTCAAATTGCTCTATACTGTATTGGACCCCTGTTAGCCCTGCAATGTGCCTACTCAT
ATATTGTACAGTCTGTTAAGTGGTATTTATAATCACCATTGTAATCTTGTGCTGCAAGAGA
GAATATTGGTGGACTGGAGATCATAGAACCTGCATGTTACCGACTTTACACAAACAGCCACAT
TTAAATTCAAAGAAAATGAGTAAGATTATAAGGTTGCCATGTGAAAACCTAGAGCATATTG
GAAATGTTCAAACCTTCTAACGTCAGATGCATTGATGACTATGTCGAATATTCTTACT
GCCATCATTATTGTTAAAGATATTGCACTTAATTGTGGAAAAATATTGCTACAATT
TTAATCTCTGAATGTAATTGATACTGTGTACATAGCAGGGAGTGATGGGGTGAACAAACTT
GGGCCAGAATATTAAACAATCATCAGGCTTTAAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSFGIMVSWKGIYFILTLFWGSFFGSIFMLSP
FLPLMFVNPSWYRWINNRLVATWLTLPMALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDWM
FLWNCLMRYSYLRLEKICLKAISLKGVPFGWAMQAAAYIFIHRKWKDDKSHFEDMIDYFCDIHEP
LQLLIFPEGTDLTENSRSNAFAEKNGLQKYEYVLHPRTGFTFVVDRLREGKNLDAVHDITVA
YPHNIPQSEKHLQGDFPREIHFHVRYPIDTLPTSKEDELQLWCHKRWEEKERLRSFYQGEKNE
YFTGQSVIPPKSELRLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG
LEIIELACYRLLHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCTGCGCCTGAGACAGCTGGCCTGACC
TCCAAATCATCCATCCACCCCTGCTGTATCTGTTTCATAGTGTGAGATCAACCCACAGGAATA
TCCATGGCTTTGTGCTCATTGGTTCTCAGTTCTACGAGCTGGTGTAGGACAGTGGCAAGT
CACTGGACCGGGCAAGTTGTCCAGGCCTGGTGGGGGAGGACGCCGTGTTCTGCTCCCTCT
TTCCTGAGACCAGTGCAGAGGCTATGGAAGTGCCTCTCAGGAATCAGTCCATGCTGTGGTC
CACCTCTACAGAGATGGGAAGACTGGGAATCTAACGAGATGCCACAGTATCAGGGAGAACTGA
GTTTGTGAAGGACTCCATTGCAGGGGGCGTGTCTCTAAGGCTAAAAAACATCACTCCCTCGG
ACATCGGCCTGTATGGGTGCTGGTTCAGTTCCAGATTACGATGAGGAGGCCACCTGGAGCTG
CGGGTGGCAGCACTGGCTCACTCCTCATTCCATCGTGGATATGTTGACGGAGGTATCCA
GTTACTCTGCCTGTCTCAGGCTGGTCCCCCAGCCCACAGCCAAGTGGAAAGGTCCACAAGGAC
AGGATTGTCTCAGACTCCAGAGCAAATGCAGATGGGTACAGCCTGTATGATGTGGAGATCTC
ATTATAGTCCAGGAAAATGCTGGAGCATATTGTGTTCCATCCACCTTGCTGAGCAGAGTCATGA
GGTGGAACTCAAAGGTATTGATAGGAGAGACGTTTCCAGCCCTCACCTGGCGCCTGGCTTCTA
TTTACTCGGGTTACTCTGTGGTGCCTGTGTGGTGTGATGGGATGATAATTGTTTCTC
AAATCCAAAGGGAAATCCAGCGGAACACTGGACTGGAGAAGAAAGCACGGACAGGCAGAATTGAG
AGACGCCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCGAAGCTCTGCG
TTTCTGATCTGAAAATGTAACCCATAGAAAAGCTCCCAGGAGGTGCCTCACTCTGAGAAGAGA
TTTACAAGGAAGAGTGTGGTGGCTCTCAGGGTTCCAAGCAGGGAGACATTACTGGGAGGTGGA
CGTGGGACAAAATGTAGGGTGGTATGTGGGAGTGTGTCGGGATGACGTAGACAGGGGAGAAACA
ATGTGACTTTGTCTCCAACAATGGGTATTGGGTCTCAGACTGACAACAGAACATTGTATTTC
ACATTCAATCCCCATTATCAGCCTCCCCCAGCACCCCTCACACGAGTAGGGTCTTCT
GGACTATGAGGGTGGGACCATCTCCTCTCAATACAAATGACCAGTCCCTTATTATACCCTGC
TGACATGTCAGTTGAAGGCTGTTGAGACCTATATCCAGCATGCGATGTGACGGAGGAAAAG
GGGACTCCCATATTCATATGTCAGTGTCTGGGATGAGACAGAGAACCCCTGCTTAAAGGC
CCCACACCACAGACCAGACACAGCCAAGGGAGAGTGCTCCGACAGGTGGCCCCAGCTCT
CCGGAGCCTGCGCACAGAGAGTCACGCCACTCTCCTTAGGGAGCTGAGGTTCTCTGCC
TGAGCCCTGAGCAGGGCAGTCACAGCTCCAGATGAGGGGGATTGGCCTGACCTGTGGAG
TCAGAAGCCATGGCTGCCCTGAAGTGGGACGGAATAGACTCACATTAGTTAGTTGTGAAA
CTCCATCCAGCTAACGATCTGAACAAGTCACAACCTCCCAGGCTCTCATTTGCTAGTCACGG
ACAGTGATTCTGCCTCACAGGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTTGAGTT
TGAGGGCACAGTGTGCTAATGATGTGTTTATATTACATTTCACCATAAACTCTGTT
TGCTTATTCCACATTAATTACTTTCTATACCAAAATCACCAGGAATAGTTATTGAACACC
TGCTTGTGAGGCTCAAAGAATAAAGAGGGAGGTAGGATTTCACTGATTCTATAAGCCCAGCAT
TACCTGATACCAAAACCAGGCAAAGAAAACAGAAGAGAGGAAGGAAAACATCAGGTCCATATCC
CTCATTAACACAGACACAAAATTCTAAATAAAATTAAACAAATTAAACTAAACAATATTTA
AAGATGATATATAACTACTCAGTGTGGTTGTCCACAAATGCAGAGTTGGTTAATATTAAAT
ATCAACCAGTGTAAATTAGCACATTAATAAGTAAAAAGAAAACCATAAAAAAAAAAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFNQFHAVVH
LYRAGEDWESKQMPQYRGRTEFKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
VAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPGQDLSSDSRANADGYSLYDVEISI
IVQENAGSILCSIHLLAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCGALCGVVMGMIIVFFK
SKGKIQAEQELDWRRKHGQAEELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF
TRKSVVASQGFQAGRHYWEVDVGQNVGWWVGVCRDDVDRGKNNVTLSNNNGYWVLRLTTEHLYFT
FNPHFISLPPSTPPTRGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLLRPIQHAMYDEEKKG
TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTCATTGCTGGTGGCAAATGATGTATCGGCCAGATGTGGTAGGGCTAGGAAAAGAG
TTTGTGGAAACCCTGGTTATCGGCCTCGTCATCTCATATCCCTGATTGTCCTGGCAGTGTGCATTGGA
CTCACTGTTCATTATGTGAGATATAATCAAAAGAAGACCTACAATTACTATAGCACATTGTCAATTACAAAC
TGACAAACTATATGCTGAGTTGGCAGAGAGGGCTCTAACAAATTACAGAAATGAGCCAGAGACTTGAAT
CAATGGTAAAAATGCATTTATAAATCTCCATTAAGGGAAGAATTGTCAAGTCTCAGGTTATCAAGTTC
AGTCAACAGAACATGGAGTGTGGCTCATATGCTGTTGATTGAGATTCACTACTGAGGATCCTGA
AACTGTAGATAAAATTGTTCAACTTGTACATGAAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTAG
ATCCTCACTCAGTTAAAATTAAAAAAATCAACAAGAACAGAACAGCTATCTAAACCATTGCTGCCGA
ACACGAAGAAGTAAAACCTAGGTCAAGTCTCAGGATCGTGGTGGACAGAAGTAGAAGAGGGTGAATG
GCCCTGGCAGGCTAGCCTGCAGTGGGATGGAGTCATCGCTGTGGAGCAACCTTAATTGCCACATGGC
TTGTGAGTGCTGCTCACTGTTACAACATATAAGAACCTGCCAGATGGACTGCTCCTTGGAGTAACA
ATAAAACCTTCGAAAATGAAACGGGTCTCCGGAGAATAATTGTCCATGAAAAATACAAACACCCATCACA
TGACTATGATATTCTCTGCAGAGCTTCTAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTGTC
TCCCTGATGCATCCTATGAGTTCAACCAGGTGATGTGATGTTGTGACAGGATTGGAGCACTGAAAAAT
GATGGTTACAGTCAAAATCATCTCGACAAGCACAGGTGACTCTCATAGACGCTACAACTTGCAATGAACC
TCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTAGAAGGAAAACAGATGCAT
GCCAGGGTGA~~CT~~GGAGGACC~~ACT~~GGTTAGTCAGATGCTAGAGATATCTGGTACCTGCTGGAATAGTG
AGCTGGGGAGATGAATGTGCGAAACCAACAAGCCTGGTGTATACAGAGTTACGGCCTTGC~~GG~~ACTG
GATTACTTCAAAACTGGTATCTAAGAGACAAAGCCTCATGGAACAGATAACATTTTTGTTTTG
GGTGTGGAGGCCATTAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTGACTGATCTCA
ATAAACTGTTGCTTGATGCATGTATTTCTCCAGCTGTCCGCACGTAAGCATTGCTGCTGCCA
GATCAACTCTGTATCTGTGAGCAATAGTTGAAACTTATGTACATAGAGAAATAGATAATACAATATTAC
ATTACAGCCTGTATTCTAGAAGTTGTCAGAATTGACTGTTGACATAAATTGTAAT
GCATATATACAATTGAAGCACTCCTTCTCAGTTCTCAGCTCCTCTCATTCAAGCAAATATCCATT
TCAAGGTGCAGAACAGGAGTGAAGAAAATATAAGAAGAAAAATCCCTACATTGACAGAA
AAGTATTAGGTGTTCTAGTGAATATTAGAAATGATCATATTCAAGGTCAAGCAAAGACA
GCAGAACATCAACTTCATCATTAGGAAGTATGGAACTAAGTTAAGGAAGTCCAGAAAGCCAAG
ATATATCCTTATTTCATTCCAAACAACTACTATGATAATGTGAAGAAGATTCTGTTTGACCT
ATAATAATTACAAACTCATGCAATGTACTTGTCAAGCAAATTAAAGCAAATTATTAACATTG
TTACTGAGGATGTCAACATATAACAATAAAATATAACCCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLIFISLIVLAVCIGLTVHYVRYNQKTYNYYSTLSFTTDKLY
AEFGREASNNFTEMSORLESMVKNAYKSPLREEFVKSQVIKFSSQQKHGVLAHMLLICRFHSTED
PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDYLNHCCGTRRSKTLGQSLRIVGG
TEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFGVTIKPSKMKRGL
RRIIVHEKYKHPHSDYDISLAEISSLSPVPTNAVHRVCLPDASYEFQPGDVMFVTGFGALKNDGYS
QNHLRQAQVTLIDATTNEPQAYNDAITPRMLCAGSLEGKTDACQGDSGGPLVSSDARDIWYLAG
IVSWGDECAKPNKPGVYTRVTALRDWITSKTGI

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAACGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAACGTTCCCTG
CCCCGATGAGCCCCCGCCGTGCGTCCCCACTATCCCCAGGCAGGGCTGGGCACCGGGCCAGC
GCCGACGATCGCTGCCGTTTGCCTTGGGAGTAGGATGTGGTGAAAGGATGGGCTCTCCCT
ACGGGGCTACAATGCCCAGAGAACGATTCCGTGAAGTGTCTGCGCTGCCTGCTACGCCCTCAA
TCTGCTCTTGGTTAATGTCCATCAGTGTGTTGGCAGTTCTGCTGGATGAGGGACTACCTAA
ATAATGTTCTCACTTAAC TGCA GAAACGAGGGTAGAGGAAGCAGTCATTGACTTACCTCCT
GTGGTCATCCGGTCATGATTGCTGTTGCTGTTCTGCATGGTACTTGGAAAGTTGCTGTCATTTCT
TGGAACGGTGAAAAGAAATCTGTTGCTTCTGCATGGTACTTGGAAAGTTGCTGTCATTTCT
GTGTAGAACTGGCTTGTGGCCTTGGACATATGAACAGGAACATTGGTTCCAGTACAATGGTCA
GATATGGTCACTTGAAAGCCAGGATGACAAATTATGGATTACCTAGATATCGGTGGCTACTCA
TGCTGGAATTTCAGAGAGAGTTAAGTGTGTTGGAGTAGTATATTCACTGACTGGTTGG
AAATGACAGAGAGATGGACTGGCCCCCAGATTCTGCTGTGTTAGAGAAATCCAGGATGTTCCAAA
CAGGCCACCAGGAAGATCTCAGTGACCTTATCAAGAGGGTTGGGAAGAAAATGTATTCCCT
TTTGAGAGGAACCAAACAACTGCAGGTGCTGAGGTTCTGGGAATCTCCATTGGGTGACACAAA
TCCTGGCCATGATTCTCACCATTACTCTGCTCTGGCTCTGTATTATGATAGAACGGGAGCCTGG
ACAGACCAAAATGATGTCCTTGAAGAACATGACAACACTCAGCACCTGTCATGTCCTCAGTAGAACT
GTTGAAACCAAGCCTGTCAAGAACATTTGAACACACATCCATGGCAAACAGCTTAATACACACT
TTGAGATGGAGGAGTTATAAAAGAAATGTCACAGAAGAAAACCACAAACTGTTTATTGGACT
TGTGAATTGGAGTACATACTATGTGTTCAGAAATATGTAGAAATAAAATGTTGCCATAAAA
TAACACCTAACATACTATTCTATGCTTAAATGAGGATGGAAAAGTTCATGTCATAAGTC
ACCACCTGGACAATAATTGATGCCCTAAAATGCTGAAGAACAGATGTCATACCCACTGTGAGCC
TGTGTATGACTTTACTGAACACAGTTATGTTGAGGCAGCATGGTTGATTGATCATTCCGCA
TCCATGCAAACGAGTCACATATGGTGGACTGGAGCCATAGTAAAGGTTGATTACTTACCAA
CTAGTATATAAAAGTACTAATTAAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTATTA
CTCAGCGATCTATTCTCTGATGCTAACATAATTATCAGAAAACATTCAATATTGGTACT
ACCTAAATGTGATTTGCTGGTTACTAAAATATTCTTACACTTAAAGAGCAAGCTAACACAT
TGTCTTAAGCTGATCAGGGATTTGTATATAAGTCTGTGTTAAATCTGTATAATTCACTGCGAT
TTCAGTTCTGATAATGTTAAGAATAACCATTATGAAAAGGAAAATTGTCCTGTATAGCATCATT
ATTTTAGCCTTCTGTTAATAAAAGCTTACTATTCTGCTCTGGCTTATATTACACATATAAC
TGTTATTAAACTAACCACTAATTGAAAATTACCACTGAGTGTGATACATAGGAATCATTATTC
AGAATGTAGTCTGGCTTTAGGAAGTATTAAGAAAATTGACATAACTTAGTTGATTGAGA
AAGGACTTGATGCTGTTCTCCAAATGAAGACTCTTTGACACTAACACTTTTAAAAA
GCTTATCTTGCCTCTCCAAACAAAGAACAGCAATAGTCTCCAAGTCATATAAAATTCTACAGAAA
TAGTGTCTTTCTCCAGAAAATGCTGTGAGAATCATTAAACATGTGACAATTAGAGATT
CTTGTGTTATTCACTGATTAATATACTGTGGCAAATTACACAGATTATTAATTTTACAA
GAGTATAGTATATTGAAATGGAAAAGTGCATTACTGTATTGTGTTATTGTTAT
TTCTCAGAATATGAAAGAAAATTAAATGTGTCATAAAATATTCTAGAGAGTAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAWSAWMRDYLNNVLTLTAETRVEEAVILTYFPVVHP
VMIAVCCFLIIVGMLGYCGTVKRNLNLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWSDMVT
LKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTEMDWPPDSCCVREFPGCSKQAHQ
EDLSDLYQEGCGKKMYSFLRGTKQLQVLRLFGISIGVTQILAMILTITLLWALYYDRREPQTDQM
MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTTATCCCACATCCTCATCCTCTTCTGATAAAGCCCCTACCAAGTGT
GATAAAAGTCTTCTCGTGAGAGCCTAGAGGCCTAAAAAGTGCTGAAAGAGAAGGGGACAAAGGAACA
CCAGTATTAAGAGGATTTCCAGTGTTCAGTGGCAGTGTCAGAAGGATGCCTCCATTCTGCTCTCACCTG
CCTCTTCATCACAGGCACCTCCGTGTCACCCGTGGCCCTAGATCCTGTTACATCAGCCTGAATGAGC
CCTGGAGGAACACTGACCACCAGTTGGATGAGTCTCAAGGTCTCTATGTGACAACCAGTGAATGGGAG
TGGTACCACTTCACGGGCATGGCGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACCACTGTGGAACCCA
CGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAGGCGACGGCATTGTGCAACGCCAGGCTTGCCAGCT
TCAATGGGAACGTGCTGTCCTGGAACACCACGGTGGAAAGTCAAGGCTTGCCCTGGAGGCTACTATGTGTATCGT
CTGACCAAGCCCAGCGTCTGCTTCCACGTCTACTGTGGTCATTTATGACATCTGCGACGAGGACTGCCATGG
CAGCTGCTCAGATAACCAGCGAGTGCACATGCGCTCCAGGAACGTGCTAGGCCCTGACAGGCAGACATGCTTG
ATGAAAATGAATGTGAGCAAAACACGGTGGCTGCAGTGAGATCTGTGTGAAACCTCAAAACTCCTACCGCTGT
GAGTGTGGGTGGCGTGTGCTAAGAAAGTGTGCAAGACTTGTGAAGACGTTGAAGGATGCCACAATAACAA
TGGTGGCTGCAGCCACTCTGCCTGGATCTGAGAAAGGCTACAGTGTGAATGTCCCCGGGCCTGGTGCTGT
CTGAGGATAACCACACTTGCCAAGTCCCTGTGTGCAAATCAAATGCCATTGAAGTGAACATCCCCAGGGAG
CTGGTTGGTGGCCTGGAGCTCTCCTGACCAACACCTCCTGCCGAGGAGTGTCCAACGGCACCCATGTCAACAT
CCTCTCTCTCAAGACATGTGGTACAGTGGTCATGTGGTAATGACAAGATTGTGCCAGCAACCTCGTGA
CAGGTCTACCCAAAGCAGACCCGGGGAGCAGCGGGACTTCATCATCCGAACCAGCAAGCTGCTGATCCCAGG
ACCTGCGAGTTCCACGCCGTACACCATTCTGAAGGATACGTTCCAACCTCGAAACTCCCCACTGGAAAT
CATGAGCCGAAATCATGGATCTTCCATTCACTCTGGAGATCTCAAGGACAATGAGTTGAAGAGCCTTAC
GGGAAGCTCTGCCACCCCAAGCTTGTGACTCCCTACTTGGCATTGAGCCGTGGTGCACGTGAGGGC
TTGGAAAGCTTGGTGGAGAGCTGCTTGCACCCCAAGATCGACGAGGTCTGAAATACTACCTCAT
CCGGGATGGCTGTGTTCAGATGACTCGTAAAGCAGTACACATCCCGGGATCACCTAGCAAAGCACTCCAGG
TCCCTGTCTCAAGTTGTGGCAAAGACCACAAGGAAGTGTGACTGCCGGTTCTGTGACTGCCGGTCTGTGGAGTG
TTGGACGAGCGTTCCCGTGTGCCAGGGTGTGCCACCCGGAATGCGTGTGGGGCAGGAGGAGGACTCAGC
CGGTCTACAGGCCAGACGCTAACAGCGGGCCGATCCGATCGACTGGGAGGACTAGTCGTAGCCATACCTC
GAGTCCCTGCATTGGACGGCTCTGCTCTTGGAGCTTCTCCCCCACCAGCCCTCTAAGAACATCTGCCAACAGC
TGGGTTCAGACTCACACTGTGAGTCAGACTCCCAGCACCAACTCACTGTGATTCTGGTCCATCAGTGGGCA
CAGGTACAGCACTGCTGAACAATGTGGCTGGGTGGGTTTCATCTTCTAGGGTTGAAAACAAACTAACTGTCCA
CCCAGAAAGACACTCACCCATTCCCTATTCTTCTAAGTAAACCTCGTGTATGGTGAATCAGAC
CACAAAATCAGAAGCTGGGTATAATATTCAAGTTACAAACCTAGAAAAATTAAACAGTTACTGAAATTATGA
CTTAAATACCAATGACTCCTAAATATGTAATTAGTTACCTTGAAATTCAATTCAAATGCAGACTAA
TTATAGGGATTGGAAGTGTATCAATAAACAGTATATAATT

FIGURE 110

MPPFLLTCLFITGTSVSPVALDPCAYISLNEPWRNTDHQLDESQGPPLCDNHVNGEWYHFTGMGDAMP
TFCIPENHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCCLWNTTVEVKACPGGYYVYRLTKPSVCFHV
YCGHFYDICDEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGCSEICVNLKNSYRCECGVGRV
LRSDGKTCEDVEGCHNNNGCSHSLGSEKGYQCECPRLVLSEDNHTCQVPVLCKSNAIEVNIPRELVGG
LELFLTNTSCRGVSNGTHVNILFSLKTCGTVVVDVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVT
CEFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFTLEIFKDNEFEYPYREALPTLKLRLDSLYFGIEPVVHV
SGLESLVESCFATPTSKIDEVLKYYLIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKFVGKDHKEVFLHCRV
LVCGVLDERSRCAQGCHRRMRRGAGGEDSAGLQGQTLTGGPIRIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTGCTCAGCGGACAAGGATGCTGGCGTGAGGGACCAAGGCCTGCCCTGCACTCGG
GCCTCCTCCAGCCAGTGCTGACCAGGGACTCTGACCTGCTGCCAGCCAGGACCTGTGTGGGGAGGCCCT
CCTGCTGCCCTGGGGTGACAATCTCAGCTCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGATGT
TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGCAAACCCGTATCCCC
ATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGT
GGTTGTCCTCATCAAGGTGATTCTGGATAAAACTACTTCCTCTGCCAGCCTCTCCACTTCATCCGA
GGAAGCAGCTGTGTGACGGAGAGCTGGACTGTCCCTGGGGAGGACGAGGAGCAGTGTCAAGAGCTTC
CCCGAAGGGCCTGCAGTGGCAGTCCGCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCAC
AGGAACTGGTTCTGCCTGTTGACAACACTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGATGG
GCTACAGCAGAGCTGTGGAGATTGGCCAGACCAAGGATCTGGATGTTGAAATCACAGAAAACAGCCAG
GAGCTTCGCATGCCGAACCTCAAGTGGCCCTGTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGCCTG
TGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGGTGGGGAGGAGGAGCCTGTGGATTCTGGCCTTGGCAGG
TCAGCATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATTGGACCCCCACTGGTCCTCACGGCA
GCCCACTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCCAGGCTCAGACAAACTGGCAG
CTTCCCACCCCTGGCTGTGCCAAGATCATCATATTGAATTCAACCCATGTACCCAAAGACAATGACA
TCGCCCTCATGAAGCTGCAGTCCCACACTTCTCAGGCACAGTCAGGCCATCTGTCTGCCCTCTT
GATGAGGAGCTCACTCCAGCCACCCACTCTGGATCATTGGATGGGCTTACGAAGCAGAATGGAGGGAA
GATGTCGACATACTGCTGCAGCGTCAGTCCAGGTATTGACAGCACACGGTCAATGCAGACGATGGT
ACCAGGGGAAGTCACCGAGAAGATGATGTGCAGGCATCCCGAAGGGGGTGTGGACACCTGCCAGGGT
GACAGTGGTGGCCCTGATGTACCAATCTGACCAGTGGCATGTGGTGGCATCGTTAGCTGGGCTATGG
CTGGGGGGCCCGAGCACCCAGGAGTATAACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCT
GGAAGGCTGAGCTGTAATGCTGCTGCCCTTGCAGTGCTGGAGGCCCTTCCCTGCCCTGCCACCT
GGGATCCCCAAAGTCAGACACAGAGCAAGAGTCCCTGGGTACACCCCTCTGCCACAGCCTCAGCAT
TTCTGGAGCAGCAAAGGCCTCAATTCTGTAAGAGACCCCTCGCAGCCAGGGGCCAGAGGAAGTCA
GCAGCCCTAGCTGCCACACTTGGTGTCCCAGCATCCCAGGGAGAGACACAGCCCAGTCAACAGGTCT
CAGGGTATTGCTAAGCCAAGGAAGGAACTTCCCACACTACTGAATGGAAGCAGGTGTCTGTAAAAGCC
CAGATCACTGTGGCTGGAGAGGAGAAGGAAAGGGTCTGCGCCAGCCCTGTCCGTCTCACCCATCCCAA
GCCTACTAGAGCAAGAAACCAGTGTAAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTTACCT
ACTGTTGTCATTGTTATTACAGCTATGCCACTATTATTAAAGAGCTGTGTAACATCTCTGGCAAAAAAA
AAAA

FIGURE 112

MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPIIIALLSLASIIIVVVLIKVILDKYFLCG
QPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSACFDN
FTEALAETACRQMGYSRAVEIGPDQDLDVVEITENSQELMRNNSGPCLSGSLVSLHCLACGKSL
KTPRVRVGEEASVDSWPWQVSIQYDKQHVCAGGSILDPHWVLTAACFRKHTDVFNWKRAGSDKL
GSFPSLAVAKIIIEFNPMPKDNDIALMKLQFPLTFSGTVRPICLPPFDEELTPATPLWIIGWG
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTQCQGDGGPLMYQS
DQWHVVGIVSWGYZGCGGPSTPGVYTKVSAYLNWIYNVWKAEL

Transmembrane domain:
amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACCTCTGGTCCAAGTGATCCACCCGCTCAGCCTCCAAAGGTGCTGTGATTA
TAGGTGTAAGCCACCGTGTCTGGCTCTGAACAACTTTTCACTAAAGCCACAGGAGT
TGAACCTGCTAGGATTCTGACTTGCTGTTGGCTAGTGCTCCTACTCCTACCTACATTAAAATC
TGTTTTTGTCTTGTAACTAGCCTTACCTCCTAACACAGAGGATCTGTCAGTGTGGCTCT
GGCCCAAACCTGACCTCACTCTGGAACGAGAACAGAGGTTCTACCCACACCGTCCCTCGAAG
CCGGGGACAGCCTCACCTGCTGGCTCTCGCTGGAGCAGTGCCCTACCAACTGTCTACGTCT
GGAGGCAGTCACTCGGGCAGTGCAGGTAGCTGAGCCTCTGGTAGCTGCGGCTTCAAGGTGGC
CTTGCCTGGCGTAGAAGGGATTTGACAAGCCGAAGATTCATAGGCGATGGCTCCACTGCC
AGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGCCAGGACGGCCGGACACCTGCTCA
GAAGCAGTGGGTGAGACATCACGCTGCCGCCATCTAACCTTTCATGTCCCTGCACATCACCTG
ATCCATGGGCTAATCTGAACTCTGTCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC
CAGAAGGGGTCTGCTTAGACCACCTGGTTATGTGACAGGACTTGCCAAATTATGGTCAGAAAAGATG
GAACGCCGGAGGAAAGCAAAGTGGCAGGGAAAGGAACCTGTGCCAAATTATGGTCAGAAAAGATG
GAGGTGTTGGTTATCACAAGGCATCGAGTCTCCTGCATTCACTGGACATGTGGGGAAAGGGCTG
CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGCCATCAGACAGCCGTTCCGCCCGAT
CCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCATGCTCTCATGCCAGGCAGCAGCCAAA
TCTGCGATCACCAGCCAGGGCAGCCGTCTGGAAAGGAGCAAGCAAAGTGACCATTCTCC
CTCCTCCCTCTGAGAGGCCCTCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGG
GGCTAATGGCTCAGTGTGGCCAGGAGGTCAAGCAAGGCCATGAGAGCTGATCAGAAGGGCTGCT
GTGCGAACACGGAAATGCCTCCAGTAAGCACAGGCTGCAAATCCCCAGGCAAAGGACTGTGTGG
CTCAATTAAATCATGTTCTAGTAATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTTGGTT
CAAATGATCTCAAGGGCCCTTATACCCAGGAGACTTGATTTGAATTGAAACCCCAAATCCA
AACCTAAGAACCAAGGTGCATTAAGAATCAGTTATTGCCGGGTGTGGCTGTAATGCCAACAT
TTGGGAGGCCAGGGGGTAGATCACCTGAGGTCAAGGAGCTGAGACAGCAGCCTGGCCAACATGG
TGAAACCCCTGTCTACTAAAAACTACAAAAACTAGCCAGGCATGGTGGTGTGCCTGTATC
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAAACCTGGAGGTGAAGGAGGCTGAGACA
GGAGAACACTTCAGCCTGAGCAACACAGCGAGACTCTGTCAGAAAAATAAAAAGAATTA
TGGTTATTTGTAA

FIGURE 114

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRD

Signal peptide:

amino acids 1-15

FIGURE 115

CAGCAGTGGTCTCTCAGCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA
AGAACCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTAAATCCAAGAAA
ATATGTAATCACTTAAGATTGTGGACTGGTGTGGTATCCTGCCCTAACTCTAATTGTCT
GTTTGGGGAGCAAGCACTCTGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT
TCTACAGCAATGGAGAGAAGAAGAAGATTACATGGAAATTGATCCTGTGACCAGAACTGAAATA
TTCAGAACGGAAATGGCACTGATGAAACATTGAAAGTGCACGACTTAAAAACGGATACTGG
CATCTACTTCGTGGTCTTCAAAATGTTTATCAAAACTCAGATTAAGTGATTCTGAATT
CTGAACCAGAACAGAGGAAATAGATGAGAATGAAGAAATTACCAACTTCTTGAACAGTCAGTG
ATTGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTCTTAAACATTCTGGA
GATTGTGATAACGTGACCATGTATTGGATCAATCCCCTCTAATATCAGTTCTGAGTTACAAG
ACTTGAGGAGGAGGGAGAAGATCTTCACTTCTGCCAACGAAAAAAAGGGATTGAACAAAAT
GAACAGTGGTGGTCCCTCAAGTGAAAGTAGAGAACCCGTACGCCAGACAAGCAAGTGAGGA
AGAACTTCCAATAATGACTATACTGAAAATGGAATAGAATTGATCCCCTGCTGGATGAGAGAG
GTTATTGTTGTATTACTGCCGTCAGGCAACCGCTATTGCCGCCGTCTGTGAACCTTACTA
GGCTACTACCCATATCCATACTGCTACCAAGGAGGACGAGTCATCTGTCGTCTGCT
TAACTGGTGGTGGCCGCATGCTGGGAGGGTCTAATAGGAGGTTGAGCTAAATGCTAAAC
TGCTGGCAACATATAATAATGCATGCTATTCAATGAATTCTGCCTATGAGGCATCTGCCCT
GGTAGCCAGCTCTCCAGAATTACTGTAGGTAATTCTCTTCATGTTCTAATAAACTTCTACA
TTATCACCAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFSKKICKSLKICGLVFGILALTLLIVLFWGSKHFPEVPKKAYDME
HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLEVHDFKNGYTGIYFVGLQKCFIKTQIKVIP
EFSEPEEEIDENEETTFFEQSVIWPAAEKPIENRDFLKNSKILEICDNVTMYWINPTLISVSE
LQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMED
ERGYCCIYCRRGNRYCRRVCEPLLGYYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCTCAGGAGCGCGTAGCTTCACACCTCGGCAGCAGGAGGGCGGCAGCTTCTCGCAGGCGCA
GGCGGGCGGCCAGGATCATGTCCACCACATGCCAAGTGGTGGCGTCTCCTGTCCATCCTGGGCT
GGCCGGCTGCATCGCGGCCACCAGGATGGACATGTGGAGCACCAGGACCTGTACGACAACCCGTACCT
CCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTCGTGAGGCAGAGTCAGGCTCACCGAATGCAGGCC
TATTCACCATCCTGGACTTCCAGCCATGCTGCAGGCAGTGCGAGCCCTGATGATCGTAGGCATCGCCT
GGGTGCCATTGGCCTCCTGGTATCCATCTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTG
CCAAAGCCAACATGACACTGACCTCCGGATCATGTTCATTGTCTCAGGTCTTGCAATTGCTGGAGTG
TCTGTGTTGCCAACATGCTGGTACTAACTTCTGGATGTCCACAGCTAACATGTACACCGCATGGTGG
GATGGTGCAGACTGTTCAGACCAGGTACACATTGGTGCAGGCTCTGTTCGTGGCTGGTCGCTGGAGGCC
TCACACTAATTGGGGTGTGATGTGCATCGCCTGCCGGGGCTGGCACCAGAAGAAACCAACTACAAA
GCCGTTCTTATCATGCCTCAGGCCACAGTGGCCTACAAGCCTGGAGGCTCAAGGCCAGCACTGGCTT
TGGGTCACACCAAAACAAGAAGATATACGATGGAGGTGCCGCACAGAGGACGAGGTACAATCTTATC
CTTCCAAGCAGCACTATGTTAATGCTCTAACGACCTCTCAGCACGGCGGAAGAAACTCCGGAGAGCTCA
CCCCAAAAACAAGGAGATCCCCTAGATTCTTGCTTTGACTCACAGCTGGAAGTTAGAAAAGCCT
CGATTCATTTGGAGAGGCCAATGGCTTAGCCTCAGTCTCTGTCTAAATATTCCACCATAAAACA
GCTGAGTTATTATGAATTAGAGGCTATAGCTCACATTTCATCCTCTATTCTTTAAATATAACT
TTCTACTCTGATGAGAGAATGTGGTTTAATCTCTCTCACATTGATGATTAGACAGACTCCCCCTC
TTCCTCCTAGTCATAAAACCCATTGATGATCTATTCCAGCTTATCCCCAGAAAACCTTGAAAGGAAA
GAGTAGACCCAAAGATGTTATTCTGCTGTTGAATTGTCTCCCCACCCCCACTTGCTAGTAATAA
ACACTTACTGAAGAAGCAATAAGAGAAAGATATTGTAATCTCTCCAGCCCAGTCTCGGTTTCTT
ACACTGTGATCTAAAGTTACCAAAACCAAGTCATTTCAGTTGAGGCAACCAACCTTCTACTGCTG
TTGACATCTCTTATTACAGCAACACCATTCTAGGAGTTCCCTGAGCTCTCCACTGGAGTCCTTTCTGT
CGCGGGTCAGAAATTGTCCTAGATGAATGAGAAAATTATTTTTAAAGTCCTAAATAGTAA
AATAAATAATGTTAGTAAATGATAACTATCTCTGTGAAATAGCCTCACCCCTACATGTGGATAGAAG
GAAATGAAAAATAATTGCTTGACATTGTCTATATGGTACTTTGTAAGTCATGCTTAAGTACAAATTCC
ATGAAAAGCTCACACCTGTAATCCTAGCAGTTGGGAGGCTGAGGAGGAAGGATCACTGAGGCCAGAAGT
TCGAGACTAGCCTGGCAACATGGAGAAGCCCTGTCTCACAAATACAGAGAGAAAAATCAGCCAGTC
TGGTGGCATAACACCTGTAAGTCCCAGCATTCCGGAGGCTGAGGTGGAGGATCACTGAGGCCAGGGAGGT
TGGGGCTGCAGTGAGCCATGATCACACCACTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAA
AATAAAAAATAATAATGGAACACAGCAAGTCCTAGGAAGTAGGTTAAAACTAATTCTTTAA

FIGURE 118

MSTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLWRSCVRQSSGFTECRP
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGL
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTGAALFVGWVAGGLTLIGGVMMCIA
CRGLAPEETNYKAVSYHASGHHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQSYP SKHDY
V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAAACTGTTCTTCTGTGGCACAGAGAACCTGCTCAAAGCAGAAGTAGCAGTCCGGAGTCC
AGCTGGCTAAACTCATCCCAGAGGATAATGGCAACCCATGCCTTAGAAATCGCTGGCTGTTCTTG
GTGGTGTGGAATGGTGGGACAGTGGCTGTCACTGTCATGCCTCAGTGGAGAGTGTGGCCTTCATT
GAAAACAACATCGTGGTTTGAAAACCTCTGGGAAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA
CATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTTCTCCGGACCTACAGGCAGCCAGAG
GACTGATGTGCTGCTCCGTGATGTCCTCTGGCTTCTGATGCCATCCTGGCATGAAATGC
ACCAGGTGCACGGGGACAATGAGAAGGTGAAGGCTCACATTGCTGACGGCTGGAATCATCTTCAT
CATCACGGGCATGGTGGTGCATCCCTGTGAGCTGGTGCATGCCATCATCAGAGATTCTATA
ACTCAATAGTGAATGTTGCCAAAAACGTGAGCTGGAGAAGCTCTACTTAGGATGGACCACGGCA
CTGGTGCTGATTGTTGGAGGAGCTGTTCTGCTGCCTTGTGCAACGAAAAGAGCAGTAGCTA
CAGATACTCGATACTTCCCATCGCACAAACCAAAAAAGTTATCACACCGGAAAGTCACCGAGCG
TCTACTCCAGAAGTCAGTATGTGTAGTTGTATGTTTTAACTTACTATAAAGCCATGCAAATG
ACAAAAATCTATATTACTTCTAAAATGGACCCAAAGAAACTTGATTACTGTTCTTAACGCCT
AATCTTAATTACAGGAACTGTGCATCAGCTATTGATTCTAAAGCTATTGAGATA
TTAAACCCAATGCTTGTAGAAAGTATAGTAATTGTTCTAAGGTGGTCAAGCATCTA
CTCTTTTATCATTACTTCTAAAATGACATTGCTAAAGACTGCATTATTACTACTGTAATTCTCC
ACGACATAGCATTATGTACATAGATGAGTGTAAACATTATCTCACATAGAGACATGCTTATATGGT
TTTATTAAAATGAAATGCCAGTCCATTACACTGAATAAATAGAACTCAACTATTGCTTCAGGGAA
ATCATGGATAGGGTTGAAGAAGGTTACTATTGTTAAAACAGCTTAGGGATTAATGTCCTCCA
TTTATAATGAAGAGTTAAACATTAAAACGCAGATATTGTCAAGGGCTTGCAATTGCT
GCTGTTTTAGCCTAGGAGTTAGAAATCCTAACTTCTTATCCTCTCCAGAGGCTTTTT
CTTGTGTATTAATTAACATTAAAACGCAGATATTGTCAAGGGCTTGCAATTGCT
TTCCAGGGCTACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAGTGTGTTAGGAAAGTG
AAAATTTGTTGTATTGAAGAAGAATGATGCATTGACAAGAAATCATATATGTATGGAT
ATATTTAATAAGTATTGAGTACAGACTTGTGAGTTCATCAATATAAATAAAGAGCAGAAAATA
TGTCTGGTTTCATTGCTTACCAAAAAACACACACACACACACACACACACACAC
GCTCCTATGTGGGTACCTGAGTCAAAATGTCATTGTTCTGTGAAAATAAATTCTTCTGTA
CCATTCTGTTAGTTACTAAAATCTGAAATACTGTATTGTTATTCAAATTGATGAA
ACTGACAATCCAATTGAAAGTTGTGTCGACGTCTGTCTAGCTAAATGAATGTGTTCTATTGCTT
TATACATTATTAATAAATTGTACATTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTAVTVMPQWRVSAFIENNIVVFENFWEGLWMNCVRQANIRMQCK
IYDSILLALSPDLQAARGLMCAASVMSFLAFMAILGMKCTRGTGNEKVKAHILLTAGIIFIITG
MVVLIPVSWVANAIIRDYNSIVNVAQKRELGEALYLGWTALVLIVGGALFCCVFCNEKSSSY
RYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCGCGCGGGTGAAGGCGATTGATGCAGCCTGCGGCGGCCTCGGAGGCCGGCGAG
CCAGACGCTGACCACGTTCTCCTCGGTCTCCTCCGCCTCCAGCTCCGCCTGCCCGCAGCC
GGGAGCCATCGACCCCAGGGCCCCGCCGCCTCCCGCAGCGGCTCCGCAGGCCCTGCTGCTCC
TGCTGCTGCAGCTGCCCGCCGTCAGCGCTCTGAGATCCCCAAGGGGAAGCAAAAGGCGCAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC
TGGTCGAGACGGGAGCCCTGGGGCCAATGTTATTCCGGTACACCTGGATCCCAGGTGGATG
GATTCAAAGGAGAAAAGGGGAATGTCAGGGAAAGCTTGAGGAGTCCTGGACACCCAAC
AAGCAGTGTTCATGGAGTTCAATTGAATTATGGCATAGATCTGGAAAATTGCGGAGTGTACATT
TACAAAGATGCGTTCAAATAGTGCCTAAAGAGTTGTTCACTGGCTCACTCGGCTAAATGCA
GAAATGCATGCTGTCAGCGTTGGTATTTCACATTCAATGGAGCTGAATGTTCAAGGACCTCTCCC
ATTGAAGCTATAATTATTGGACCAAGGAAGCCCTGAAATGAATTCAACAATTAAATTATCG
CACTTCTCTGGAAGGACTTGTGAAGGAATTGGTGGATTAGTGGATGTTGCTATCTGG
TTGGCACTTGTTCAGATTACCAAAAGGAGATGCTTCACTGGATGGAATTCAAGGTTCTCGC
ATTATTGAAGAACTACCAAAATAAATGCTTAATTTCATTGCTACCTCTTTTATTATGCC
TTGGAATGGTCACTAAATGACATTAAATAAGTTATGTATACATCTGAATGAAAAGCAAAG
CTAAATATGTTACAGACCAAAGTGTGATTTCACACTGTTAAATCTAGCATTTCATTG
CTTCAATCAAAAGTGGTTCAATATTTAGTTGGTAGAATACTTCTTCATAGTCACATT
CTCTCAACCTATAATTGGAATTGTTGCTGGCTTTGTTCTCTTAGTATAGCATT
AAAAAAATAAAAAGCTACCAATCTTGACATTGAAATGTTAAGAATTTTTATATCTGT
TAAATAAAATTATTCCAACA

FIGURE 122

MRPQGPAAASPQRLRGLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAGVPGR
DGSPGANVIPGTPGIPGRDGFKGEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK
MRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECSGPLPIEAIYLDQGSPEMNSTINIHRTS
SVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCGCGTACGGGCTCTCCTGCCTCTGGGCTCCAACGCAGCTCTGTGGCTGAA
CTGGGTGCTCATCACGGAACTGCTGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA
ATTGCCTGGAAGAATAACATCATGTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTTTTA
ACCGCCCTCCCCACCCCCAAAAAACTGTAAGATGCAAAACGTAATATCCATGAAGATCC
TATTACCTAGGAAGATTGATGTTGCTGCGAATGCGGTGTTGGATTATTGTTCTGGAG
TGTTCTGCGTGGCTGGCAAAGAATAATGTTCAAATCGGTCCATCTCCAAGGGTCCAATT
TCTTCCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGCTGTCACTG
GCCCTAAGCCAAGCAAAGACCTAAGGACGACCTTGAACAATAACAAAGGATGGTTCAATG
TAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTAGCCCCACTGTCTTACTGACAATG
CTTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTGTAGGTGTAGGGCAAATGGTATATTGTGA
ATCTCAGAAATTACAGGAGATACCCCTAAGTATATCTGCTGGTGCTTAGGTTGTCCCTCGCT
ATAACAGCCTCAAAAACCTTAAGTATAATCAATTAAAGGGCTAACCGAGCTCACCTGGCTATAC
CTTGACCATAACCATACTGCAATATTGACGAAAATGCTTTAATGGAATACGCAGACTCAAAGA
GCTGATTCTTAGTCCAATAGAATCTCCTATTTCTTAACAATACCTCAGACCTGTGACAAATT
TACGGAACTTGGATCTGCTCTATAATCAGCTGCATTCTCTGGGATCTGAACAGTTGGGGCTTG
CGGAAGCTGCTGAGTTACATTACGGCTAACCTCCCTGAGAACCATCCCTGTGCGAATATTCCA
AGACTGCCGCAACCTGGAACCTTGGACCTGGGATAAACGGATCCGAAGTTAGCCAGGAATG
TCTTGCTGGCATGACTCAAAGAACCTCACCTGGAGCACAACTCAATTCCAGCTCAAC
CTGGCCCTTTCCAAGGGTGGTCAGCCTCAGAACCTTACTTGCAAGGAAATTCAGTGT
CATAGGACAGACCATGTCCTGGACCTGGAGCTCCTAACAAAGGCTGATTATCAGGCAATGAGA
TCGAAGCTTCAGTGGACCCAGTGTGTTCCAGTGTGTCCTGAACGGCTCAACCTGGAT
TCCAACAAGCTCACATTATTGGTCAAGAGATTGGATCTGGATATCCCTCAATGACATCAG
TCTTGCTGGGAATATGGGAATGCAGCAGAAATATTGCTCCCTGTAAACTGGCTGAAAGTT
TTAAAGGTCTAAGGGAGAATACAATTATCTGTCAGTCCAAAGAGCTGCAAGGAGTAAATGTG
ATCGATGCAGTGAAGAACTACAGCATCTGTGGCAAAGTACTACAGAGAGGTTGATCTGGCAG
GGCTCTCCAAAGCCGACGTTAAGCCCAAGCTCCCCAGGCCAACAGAGCCGGCAGAGACCATCT
TGCCCCCGACGGTGGAGCCACAGAGCCGGCAGAGACCGATGCTGACGCCGAGCACATCT
TTCCATAAAATCATCGCGGGCAGCGTGGCCTTCTGTCCGTGCTCGTCATCCTGCTGGTTAT
CTACGTGTCAAGGAAGCGGTACCCCTGCGAGCATGAAGCAGCTGCAGCAGCGCTCCCTCATCGAA
GGCACAGGAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCCAGCACCCAGGAATTTATGTA
GATTATAAACCCACCAACACGGAGACCAGCGAGATGCTGTAATGGACGGGACCCCTGCACCTA
TAACAAATCGGGCTCCAGGGAGTGTGAGGTATGAAACCATGTTGATAAAAGAGCTCTAAAGCT
GGGAAATAAGTGGTGGTTATTGAACTCTGGTGAATCAAGGGAACGCGATGCCCTCCCC
TTCCCTCTCCCTCTACTTGGTGGCAAGATCCTCTGTCCGTTAGTGCATTATAACT
GGTCATTTCTCTCATACATAATCAACCCATTGAAATTAAATACCACAAATCAATGTGAAGCTT
GAACCTCCGGTTAATATAATACCTATTGTATAAGACCCCTTACTGATTCCATTAATGTCGCA
GTTTAAGATAAAACTTCTTCATAGGTAAAAAAAAAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPTVLLTMLSSAERGCPKGRCRCEGMVYCESQKLQEIPSSISAGCLG
LSLRYNSLQKLKYNQFKGLNQLTWLYLDHNNHISNIDENAFNGIRRLKELILSSNRISYFLNNNTFR
PVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHLRSNSLRTIPVRIFQDCRNLELLDLGYNRIRS
LARNVFAGMIRLKELHLEHNQFSKLNLAFLPRLVSLQONLYLQWNKISVIGQTMSTWSSLQRLDL
SGNEIEAFSGPSVFQCVNLQRLNLDNSNKLFIGQEILDWISLNDISLAGNIWECSRNICSIVN
WLKSFKGLRENTIICASPKELOQGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE
SKPPLPPTVGATEPGPETDADAEHISFKHIIAGSVALFLSVLVLVLLVIYVSWKRYPASMKQLQQR
SLMRRHRKKKRQSLKQMTPSTQEFYVDYKPTNTETSEMLLNGTGPCTYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTGCCTACTGCTGAATGTCGCCGGAGGAGGGCTTGCCGCTG
ACCCAGAGATGGCCCCGAGCGAGCAAATCCTACTGTCCGGCTGCGCGCTACCGTGGCCGAGCT
AGCAACCTTCCCTGGATCTCACAAAACCGACTCCAATGCAAGGAGAACGAGCTTGCTC
GGTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAGGGGAATGGTGCACAGCCCTAGGGATC
ATTGAAGAGGAAGGCTTCTAAAGCTTGGCAAGGAGTGACACCCGCCATTACAGACACGTAGT
GTATTCTGGAGGTGCAATGGTCACATATGAACATCTCCGAGAGGTTGTGTTGGCAAAAGTGAAG
ATGAGCATTATCCCCTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTATTGCCAGTT
TTAGCCAATCCAACGTGACCTAGTGAAGGTTAGATGCAATGGAAGGAAAAGGAAACTGGAAGG
AAAACCATTGCGATTCTGTGGTACATCATGCATTGCAAAATCTAGCTGAAGGAGGAATAC
GAGGGCTTGGCAGGCTGGTACCCAATATAACAGAGCAGCACTGGTGAATATGGGAGATTAA
ACCACTTATGATACTGAAACACTACTTGGTATTGAATAACCACTTGAGGACAATATCATGAC
TCACGGTTATCAAGTTATGTTCTGGACTGGTAGCTCTATTCTGGAACACCAGCCATGTCA
TCAAAAGCAGAATAATGAATCAACCACGAGATAACAAGGAAGGGACTTTGTATAAATCATCG
ACTGACTGCTTGATTCAAGGTGAAGGATTGAGTCTATATAAAGGCTTTTAC
ATCTGGCTGAGAATGACCCCTGGTCAATGGTGTCTGGCTTACTTATGAAAAAATCAGAGAGA
TGAGTGGAGTCAGTCCATTTAA

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMQGEAALARLGDGARES
APYRGMVRTALGIIEEEGFLKLWQGVTPAIYRHVVYSGGRMVTYEHREVVFGKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMOMEGRKRKLEGKPLRFRGVHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCGLVASILGTPADVIKSRI
MNQP
RDKQGRGLLYKSSTDCLIQAVQGE~~G~~FMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 127

CGCGGATCGGACCCAAGCAGGTGGCGGCCGGCAGGAGAGCGCCGGCGTCAGCTCCTCGAC
CCCCGTGTGGGCTAGTCCAGCGAGGCGGACGGCGGCCGTGGGCCATGGCCAGGCCGGCATGG
AGCGGTGGCGCGACCGGCTGGCGCTGGTACGGGGCCTCGGGGGCATGGCGGCCGTGGCC
CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGCTGCGCCGCACTGTGGCAACATCGAGGA
GCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCGGGACTTGTACAGATGTGACCTAT
CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTAGACATC
TGCATCAACAATGCTGGCTTGGCCCGGCTGACACCCCTGCTCTCAGGCAGCACAGTGGTTGGAA
GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAAGTCCATGA
AGGAGCGGAATGTGGACGATGGCACATCATTAAACATCAATAGCATGTCTGGCCACCGAGTGT
CCCCTGTCTGTGACCCACTTCTATAGTGCACCAAGTATGCCGTACTGCCTGACAGAGGGACT
GAGGCAAGAGCTTCGGGAGGCCAGACCCACATCCGAGGCCACGTGCATCTCCAGGTGTGGTGG
AGACACAATTGCCTTCAAACCTCCACGACAAGGACCTGAGAAGGCAGTGCACCTATGAGCAA
ATGAAGTGTCTCAAACCCGAGGATGTGGCGAGGCTGTTATCTACGTCTCAGCACCCCCGCACA
CATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACCTAGTGACTGTGGAGCTCC
TCCTTCCCTCCCCACCCCTCATGGCTTGCCTCCTGCCTGGATTAGGTGTTGATTCTGGAT
CACGGGATACCACCTCCTGTCCACACCCGACCAGGGCTAGAAAATTGTTGAGATTTTATA
TCATCTTGTCAAATTGCTTCAGTTGAAATGTGAAAAATGGCTGGGAAAGGAGGTGGTGTCCC
TAATTGTTTACTTGTAACTTGTCTGTGCCCTGGCACTGGCCTTGTCTGCTCTCAGTG
TCTTCCCTTGACATGGAAAGGAGTTGTGGCCAAATCCCCATCTTGCACCTCAACGTCTG
TGGCTCAGGGCTGGGTGGCAGAGGGAGGCCTCACCTTATATCTGTGTTATCCAGGGCTCC
AGACTCCTCCTGCCTGCCCACTGCACCCCTCCCCCTATCTATCTCCTCTCGGCTCCCC
AGCCCAGTCTGGCTTGTCCCTGGGTCACTCCACTGTACTCTGACTATGGCAG
CAGAACACCAGGGCTGGCCAGTGGATTGATCATTAAAAAGAAAAATCGCAACCAA
AAAAAAAAAA

FIGURE 128

MARPGMERWRDRRLALVTGASGGIGAAVARALVQQGLKVVGCARTVGNIEELAAECKSAGYPGTLI
PYRCDLSNEEDILSMFSAIRSQHSGVDICINNAGLARPDTLLSGSTSGWKDMFNVLALSICTR
EAYQSMKERNVDDGHIININMSGHRVLPLSVTHFYSATKYAVTALTEGLRQELREAQTHIRATC
ISPGVVETQFAFKLHDKDPEKAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AACTCTACATGGGCCTCCTGCTGGTGTCTTCCTCAGCCTCCTGCCGGTGGCTACACCAT
CATGTCCTCCCACCCCTCTTGACTGCGGCCGTTCAGGTGCAGAGTCTCAGTTGCCCGGGAGC
ACCTCCCCTCCCGAGGCAGTCTGCTCAGAGGCCCTGGCCCAGAATTCCAGTTCTGGTTCATGC
CAGCCTGTAAAAGGCCATGGAACCTTGGGTGAATCACCGATGCCATTAAAGAGGGTTCTGCCA
GGATGGAAATGTTAGGTCGTTCTGTCTGCGCTGTTCAATTTCAGTAGCCACCAGCACCTGTGG
CCGTTGAGTGCTTGAAATGAGGAACTGAGAAAATTAAATTCTCATGTATTTCTCATTTATTAA
TTAATTAACTGATAGTTGTACATATTGGGGTACATGTGATATTGGATACTGTATACAA
TATATAATGATCAAATCAGGGTAACACTGGGATATCCATCACATCAAACATTATTTTATTCTT
TTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGCCATCTCAGCTTACTGCAAC
CTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCTCCACCTCCAAAGTAGCTGGACTACAGGCAT
GCACCACAATGCCCAACTAATTGTATTTTAGTAGAGACGGGGTTGCCATGTTGCCAGG
CTGGCCTGAACCTCCTGGCCTCAAACAATCCACCTGCCTCGGCCTCCAAAGTGTATGATTACA
GGCGTAGGCCACCGTGCCTGGCTAACATTTATCTTTCTTGTGTTGGAACTTGAAATTAT
ACAATGAATTATTGTTAACTGTCATCTCCCTGCTGTGCTATGGAACACTGGGACTTCTCCCTCT
ATCTAACTGTATTTGTACCAAGTTAACCAACCGTACTTCATCCCCACTCCTCTATCCTTCCC
AACCTCTGATCACCTCATTCTACTCTACCTCCATGAGATCCACTTTAGCTCCACATGTG
AGTAAGAAAATGCAATATTGTCTTCTGTGCCCTGGCTTATTCACTAACATAATGACTCCTG
TTCCATCCATGTTGCTGCAAATGACAGGATTGTTCTTAATTCAATTAAAATAACCACACATG
GCAAAAA

FIGURE 130

MGLLLLVLFLSLLPVAYTIMSLPPSFDGPFRCRVSVAREHLP SRGSLLRGPRPRI PVLVSCQPV
KGHGTLGESPM PFKRVFCQDG NVRSFCVCAVHFSSHQPPV AVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTGTATAAAGACCTAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
ATCTTCCTCATCGGGACTAAAATTGGGCTGTTCTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCTG
TCCATCTGTGTGTCGCTGCGATGCGGGTTTCATTACTGTAATGATCGCTTCTGACATCCATTCCAACAG
GAATACCAGAGGATGCTACAACTCTACCTTCAGAACAAACAAATAATGCTGGATTCTTCAGAT
TTGAAAAACTTGCTGAAAGTAGAAAGAATATACCTATACCACAAACAGTTAGATGAATTCTACCAACCT
CCCAAAGTATGTAAAAGAGTTACATTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTCAA
AAATTCCCTATCTGGAAGAATTACATTAGATGACAACCTGTCTCTGCAGTTAGCATAGAAGAGGGAGCA
TTCCGAGACAGCAACTATCTCCGACTGCTTCTGCCCCTGTAATCACCTAGCACAATTCCCTGGGTTT
GCCAGGACTATAGAAGAACTACGCTGGATGATAATCGCATATCCACTATTCATCACCCTCTTCAAG
GTCTCACTAGTCTAAAACGCCTGGTTCTAGATGGAAACCTGTTGAACAATCATGGTTAGGTGACAAAGTT
TTCTCAACCTAGTTAATTGACAGAGCTGTCCTGGTGCAGAATTCCCTGACTGCTGCACCAGTAAACCT
TCCAGGCACAAACCTGAGGAAGCTTATCTCAAGATAACCACATCAATGGGTGCCCAATGCTTTT
CTTATCTAAGGCAGCTATCGACTGGATATGTCCAATAACCTAAAGTAATTACCTCAGGGTATCTT
GATGATTGGACAATATAACACAACGTGATTCTCGAACAAATCCCTGGTATTGCGGGTGCAAGATGAAATG
GGTACGTGACTGGTTACAATCACTACCTGTGAAGGTCAACGTGCGTGGCTCATGTGCCAAGCCCCAGAAA
AGGTTCGTGGGATGGCTATTAAGGATCTCAATGCAGAACTGTTGATTGTAAGGACAGTGGGATTGTAAGC
ACCATTCAAGATAACCACGTCAATACCCAAACACAGTGTATCCTGCCAAGGACAGTGGCCAGCTCCAGTGAC
CAAACAGCCAGATATTAAGAACCCAAAGCTCACTAAGGATCAACAAACACAGGGAGTCCCTCAAGAAAAA
CAATTACAATTACTGTGAAGTCTGCACCTCTGATACCATTCAATCTCTGGAAACTTGCTCTACCTATG
ACTGCTTGAGACTCAGCTGGCTAAACTGGGCCATAGCCGGCATTTGGATCTATAACAGAAACAATTGT
AACAGGGAACCCAGTGAAGTACTGGTACAGCCCTGGAGCCTGATTCCCTATAAAAGTATGCATGGTC
CCATGGAAACCCAGCAACCTCTACCTATTGATGAAACTCCTGTTGATTGAGACTGAAACTGCACCCCT
CGAATGTACAACCCCTACAACCACCCCTCAATCGAGAGCAAGAGAAAAGAACCTTACAAAACCCAAATTAC
TTTGGCTGCCATCATTGGTGGGCTGTGGCCCTGGTACCAATTGCCCTCTTGCTTGTGTTGGTATG
TTCATAGGAATGGATCGCTTCTCAAGGAACGTGCAATAGCAAAGGGAGGAGAAGAAAGGATGACTAT
GCAGAAGCTGGCACTAAGAAGGACAACCTATCCTGAAATCAGGAAACTTCTTCAGATGTTACCAAT
AAGCAATGAACCCATCTCGAAGGAGGAGTTGTAATACACACCATAATTCCCTCTAATGGAATGAATCTGT
ACAAAAACAATCACAGTGAAGCAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC
TCACACTCATGATGCTGAAGGACTCACAGCAGACTTGTGTTGGTTTTAACCTAAGGGAGGTGATG
GT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVRCDAAGFIYCNDRFLTSIPTGIPEDATTLYL
QNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIRTITYDSLSKIPYL
EELHLDDNSVSAVSIEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSL
QGLTSLKRLVLDGNLLNNHGLGDKVFFNLVNLTELSVRNSLTAAAPVNLPGTNLRKLYLQDNHIN
RVPPNAFSYLRQLYRLDMSNNNLSNLPQGIFDDLDNITQLILRNNPWYCGCKMKWVRDWLQSLPV
KVNRVGLMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQITTAIPNTVYPAQGQWPAPVTKQPD
IKNPKLTKDQQTGSPSRKTITITVKSVTSDEHISWKLALPMTALRLSWLKLGHSPAEGSITET
IVTGERSEYLVTALEPDSPYKVCVPMETSNLYLFDETPVCETETAPLRMYNPTTLNREQEKE
PYKNPNLPLAAIIGGAVALVTIALLALVCWYVHRNGSLFSRNCAYSKGRRRKDDYAEAGTKKD
ILEIRETSFQMLPISNEPISKEEFVVIHTIFPPNGMNLYKNNHSESSSNRSYRDSGIPDSDHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCACTCCCCCTGCAGCCACCCCTTCCCAGAGTCCTTGCCTCAGGCCACCCAGGCTTCTGGCA
GCCCTGCCGGGCCACTTGTCTTCATGTCTGCCAGGGGAGGTGGGAAGGAGGTGGGAGGGAGGGCG
TGCAGAGGCAGTCTGGGCTTGGCCAGAGCTCAGGGTGTGAGCGTGTGACCAGCAGTGAGCAGAG
GCCGGCCATGCCAGCCTGGGCTGCTGCTCCTGCTCTTACTGACAGCACTGCCACCCTGTGGT
CCTCCTCACTGCCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCTGATCCTGTCT
GCGCTGGAGAGAGGCCACCGTCTTCCTAGAACACAGAGGCTGCCTGAAATCAACCTGGATGGCATGGT
GGGGGTCCGAGTGCTGGAAGAGCAGCTAAAAAGTGTCCGGGAGAAGTGGGCCCAGGAGCCCTGC
TGCAGCCGCTGAGCCTGCGCGTGGGATGCTGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
CTCCACTACCTCAAGCTGAGTGATCCCAAGTACCTAACAGAGAGTTCCAGCTGACCCTCCAGCCGG
GTTTGGAAAGCTCCCACATGCCCTGGATCCACACTGATGCCCTGGTGTACCCCACGTTGGGC
CCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCCCTGGTGCAGCTGCTGGAAACCGGGACG
GACAGCAGCGAGCCCTGCGGCCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCGGCTGCTC
AGGCTACTGCCGTCCCACCAACTGCTCTTCCCTGGGCCAGAATGAGGGGATGCACACAGG
GACCACTCCAACAGAGCCAGGACTATATCAACCTCTGCGCCAACATGATGGACTTGAACCGC
AGAGCTGAGGCCATCGGATACGCCTACCCCTACCCGGACATCTTCATGGAAAACATCATGTTCTG
TGGAATGGCGGCTTCTCCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA
AACAGCAGGAAGGATGCTCAGGAGCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA
TATCAGCAGCATTTCGAGGAGAGTGAAGAGGGGAGAAAAACAATTCCAGATTCTCGCTCTGT
TGCTCAGGCTGGAGTACAGTGGCGCAATCTGGCTCACTGCAACCTTGCCTGGTTCAAGC
AATTCTCTGCCTCATCCTCCCGAGTAGCTGGACTACAGGAGCGTGCACCACACCTGGCTAAT
TTTATATTTTTAGTAGAGACAGGGTTCATCATGTTGCTCATGCTGGTCTCGAACTCCTGAT
CTCAAGAGATCCGCCACCTCAGGCTCCAAAGTGTGGATTATAGGTGTGAGCCACCGTGTCTG
GCTGAAAAGCACTTCAAAGAGACTGTGTTGAATAAAGGGCAAGGTTCTGCCACCCAGCACTC
ATGGGGCTCTCCCTAGATGGCTGCTCCTCCCACAAACACAGCCACAGCAGTGGCAGCCCTGG
GTGGCTTCTATACATCCTGGCAGAATACCCCCCAGCAAACAGAGAGCCACACCCATCCACACCG
CCACCAAGCAGCCGCTGAGACGGACGGTTCCATGCCAGCTGCCCTGGAGGGAGGAACAGACCC
TTTAGTCCTCATCCCTTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG
ATAAGCAAAGCCACCCGACACCCAAATCTTGGAAAGCCCTGAGTAGGCAGGGCAGGGTAGGTGGG
GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAGTCAACTGCAACTGAAAAAA

FIGURE 134

MSARGRWEGGGRRACRGSGLARAQGAERVTSSEQRPMASLGLLLLLTLAPPLWSSSLPGLD
TAESKATIADLILSALERATVFLEQRLPEINLDGMVGVRVLEEQLKSVREKWAQEPLLQPLSLRV
GMLGEKLEAAIQRSLHYLKLSDPKYLREFQTLQPGFWKLPHAWIHTDASLVYPTFGPQDSFSEE
RSDVCLVQLLGTGTDSEPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQGPLQQSQD
YINLFCANMMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILSWQKQQEGCFG
EPDAEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFQFSCLILP
SSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRDPPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCATGGCGGCCGCTCTGGGGCTTCTTCCCCTGCTGCTG
CTGCTGCTATCGGGGATGTCCAGAGCTCGGAGGTGCCGGGCTGCTGCTGAGGGATCGGGAGG
GAGTGGGTGGCATAGGAGATCGCTCAAGATTGAGGGCGTGCAGTTGTCCAGGGTGAAGC
CTCAGGACTGGATCTGGCGGCCGAGTGCTGGTAGACGGAGAAGAGCACGTCGGTTCTTAAG
ACAGATGGAGTTTGTGGTTATGATATACTTCTGGATCTTATGAGTGAAGTTGTATCTCC
AGCTTACAGATTTGATCCCCTCGAGTGGATATCACTCGAAAGGAAAATGAGAGCAAGATATG
TGAATTACATCAAAACATCAGAGGTTGTCAAGACTGCCCTATCCTCTCCAAATGAAATCTCAGGT
CCACCTTCTTACTTTATTAAAAGGAATCGTGGGCTGGACAGACTTCTAATGAACCCAATGGT
TATGATGATGGTCTTCCTTATTGATATTGTGCTTCTGCCTAAAGTGGTCAACACAAGTGATC
CTGACATGAGACGGAAATGGAGCAGTCAATGAATATGCTGAATTCCAACCAGTGCCTGAT
GTTTCTGAGPTCATGACAAGACTCTCTTCAAAATCATCTGGCAAATCTAGCAGCGGCAGCAG
TAAAACAGGCAAAAGTGGGCTGGCAAAAGGAGGTAGTCAGGCCGTCCAGAGCTGGCATTGCAC
AAACACGGCAACACTGGTGGCATCCAAGTCTGGAAAACCAGTGAAGCAACTACTATAAATT
GAGTCATCCCGACGTTGATCTTACAACGTGTATGTT
AACTTTTAGCACATGTTGTACTTGGTACACGAGAAAACCCAGCTTCATTTGTCTGTAT
GAGGTCAATATTGATGTCACTGAATTAATTACAGTGTCTATAGAAAATGCCATTAAATAAATTAT
ATGAACTACTATACTATTATGATATTAAACATCTAATCCAGAAATCAAAAAAAA
AAAAAAA

FIGURE 136

MAAALWGFFPVLLLLLSDVQSEVPGAAAEGSGGSGVGIGDRFKIEGRAVPGVKPQDWISAA
RVLVDGEEHVGFLKTDGSFVVHDIPSGSYVVEVVS PAYRFDPVRVDITSKGKMRARYVNYIKTSE
VVRLPYPLQMKSSGPPSYFIKRESWGWTDFLMNPVMMVLPLLI FVLLPKVVNTSDPDMRREME
QSMNMLNSNHELPDVSEFMTRLFSSKSSGKSSGSSKTGSGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTCTGTGAGATTGATTCGATTCCTCCCCAGTCCCCGTGGGCTGAGGGGA
CCAGAAGGGTGAGCTACGTTGGCTTCTGGAAGGGGAGGCTATATCGTCAATTCCCCAAACAA
GTTTGACATTCCCCCTGAAATGTCATTCTATCTATTCACTGCAAGTGCTGCTGTTCCAGGC
CTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACGACCTGTGC
CACCAACTCGCACTCAGACTCTGAACCTCAGACCTGAAATCTCTTCACGGGAGGCTTGGCAGT
TTTCTTACTCCTGTGGTCTCCAGATTCAGGCCTAACGATGAAAGCCTCTAGTCTTGCCCTCAGC
CTTCTCTGCTGCGTTTATCTCCTATGGACTCCTCCACTGGACTGAAGACACTCAATTGGG
AAGCTGTGTGATGCCACAAACCTTCAGGAAATACGAAATGGATTCTGAGATACAGGGCAGTG
TGCAAGCCAAAGATGGAAACATTGACATCAGAACTTAAGGAGGACTGAGTCTTGCAAGACACA
AAGCCTGCGAATCGATGCTGCCCTCGCCATTGCTAACGACTCTATCTGGACAGGGTATTAA
AAACTACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGCAGCCTGCCATTGAGGAGGAAGCAATG
CCATCAAGAAGGACCTCCGGCTCTCATGCCACATGACATGCCATTGTGGGAGGAAGCAATG
AAGAAATACAGCCAGATTCTGAGTCACTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC
TTTGGGGAACTAGACATTCTCTGCAATGGATGGAGGAGACAGAATAGGAGGAAAGTGATGCTG
CTGCTAACGAAATATTGAGGTCAAGAGCTCCAGTCTCAATACCTGCAGAGGAGGCATGACCCAA
ACCACCATCTTTACTGTACTAGTCTGTGCTGGTCACAGTGTATCTTATGCATTACTTG
CTTCCTGCATGATTGTCTTATGCATCCCCAATCTTAATTGAGACCATACTGTATAAGATT
TGTAATATCTTCTGCTATTGGATATATTATTAGTTAATATATTATTATTGCTATT
ATGTATTATTTTACTGGACATGAAACTTAAAAAAATTCAACAGATTATTTATAACCTG
ACTAGAGCAGGTGATGTATTACAGTAAAAAAACCTGTAAATTCTAGAAGAGTGG
CTAGGGGGTTATTCAACTAAGGACATATTACTCATGCTGATGCTGTGAGAT
ATTGAAATTGAACCAATGACTACTTAGGATGGGTTGTGGAATAAGTTGATGTGGAATTGCAC
ATCTACCTTACAATTACTGACCACCCAGTAGACTCCCCAGTCCATAATTGTGTATCTCCAG
CCAGGAATCCTACACGGCCAGCATGTATTCTACAAATAAAGTTCTTGCAACCAAAAAAA
AAAAAA

FIGURE 138

MRQFPKTSFDISPEMSFSIYSLQVPAVPGLTCWALTAEPGWQNKGATTCATNSHSDSELRPEIF
SSREAWQFFLLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
FSEIRGSVQAKDGNIDIRILRRTESLQDTK PANRCCLLRHLLRLYLDRVFKNYQTPDHYTLRKIS
SLANSFLTIKKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEET
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGTCGGTCCGCATCCAGCC
TAGCGTGTCCACG**A**TGCGGCTGGGCTCCGGACTTCGCTACCTGTTGCGTAGCGATCGAGGTGC
TAGGGATCGCGGTCTTCCTCGGGATTCTTCCCGCTCCGTCGTTCTGCCAGAGCGGAA
CACGGAGCGGAGCCCCAGCGCCGAACCTCGCTGGAGCCAGTCTAACTGGACCACGCTGCC
ACCACCTCTTCAGTAAAGTTGTTAGTGTCTGATAGATGCCTTGAGAGATGATTTGTGTTG
GGTCAAAGGGTGTGAAATTATGCCCTACACAACCTACCTTGTGGAAAAAGGAGCATCTCACAGT
TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCTCGAATCAAGGCATTGATGACGGGGAG
CCTTCCTGGCTTGTGACGTACAGGAACCTCAATTCTCCTGCACGTGCTGGAAAGACAGTGTGA
TAAGACAAGCAAAGCAGCTGGAAAAAGAATAGTCTTTATGGAGATGAAACCTGGGTTAAATTA
TTCCCAAAGCATTGTGGAATATGATGGAACACCTCATTTCGTGTCAGATTACACAGAGGT
GGATAATAATGTCACGAGGCATTGGATAAGTATTAAAAGAGGAGATTGGGACATATTAATCC
TCCACTACCTGGGGCTGGACCACATTGCCACATTTCAGGGCCAACAGCCCCCTGATTGGCAG
AAGCTGAGCGAGATGGACAGCGTGTGATGAAGATCCACACCTCACTGCAGTCGAAGGAGAGA
GACGCCTTACCAATTGCTGGTCTTGTTGACCATGGCATGTCTGAAACAGGAAGTCACG
GGGCCTCCTCCACCGAGGGAGGTGAATACACCTCTGATTAAATCAGTTCTGCTGGAAAGGAAA
CCCAGGTGATATCCGACATCCAAAGCACGTCCAA**T**AGACGGATGTGGCTGCGACACTGGCGATAGC
ACTTGGCTTACCGATTCCAAAAGACAGTGTAGGGAGCCTCTATTCCAGTTGTGGAAGGAAGAC
CAATGAGAGAGCAGTTGAGATTTCATTTGAATACAGTGCAGCTAGTAAACTGTGCAAGAG
AATGTGCCGTATATGAAAAGATCCTGGGTTGAGCAGTTAAAATGTCAAGAAAGATTGCATGG
GAACGGATCAGACTGTACTGGAGGAAAAGCATTCAAGTCCTATTCAACCTGGGCTCCAAGG
TTCTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTGTCCCTGAGTCACAAGTGGCCAG
TTCTCACCCCTGCTCTGCTCAGCGTCCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAAGTCCA
CTGTCATCTCCTGGGTTCTGCTCTTTATTGGTGAATCCTGGTTCTTCGGCCGTTCACGT
CATTGTGTCACCTCAGCTGAAAGTTGCTACTTCTGTCAGCTGCTGGCTGGCGAGGCT
GCCTTCGTTACCAAGACTCTGGTGAACACACTGGTGTGCAAGTGCAGTGCAGTGCCTGGAC
AGGGGCCTCAGGGAAGGACGTGGAGCAGCCTTATCCAGGCCTCTGGGTGTCCCACAGGTG
TTCACATCTGTGCTGTCAGGTCAAGTCCTCAGTTCTGAAAGCTAGGTTCTGCACTGTTAC
CAAGGTGATTGTAAAGAGCTGGCGGTACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGAA
TCGGACAGCCTCCCAGCAGAGGTGTGGAGCTGAGCTGAGGGAAAGAAGAGACAATCGGCCTGGA
CACTCAGGAGGGTCAAAGGAGACTTGGTCGCACCACTCATCCTGCCACCCCCAGAATGCATCCT
GCCTCATCAGGTCCAGATTCTTCAAGGCGGACGTTCTGTTGAAATTCTTAGTCCTGGCC
TCGGACACCTTCATTGTTAGCTGGGAGTGGTGGTGAAGGCAGTGAAGAAGAGGCGGATGGTCAC
ACTCAGATCCACAGAGCCCAGGATCAAGGGACCCACTGCAGTGGCAGCAGGACTGTTGGCCCC
ACCCCAACCTGCACAGCCCTATCCCCCTTGCTTGAGCCGTAGAGGCCGTGCTGAGTGT
CTGACCGAGACACTCACAGCTTGTCATCAGGGCACAGGCTCTCGGAGCAGGATGATCTGTG
CCACGCTTGCACCTCGGGCCATCTGGCTCATGCTCTCCTGCTATTGAATTAGTACCTAG
CTGCACACAGTATGTAGTTACAAAAGAATAAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEHGAEPPEPEPSAGASSNWTLLPPPLF
SKVVIVLIDALRDDFVFGSKGVKFMPYTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDSVIRQAKAAGKRIVFYGDETWVKLFPKHFVEYDGTSFFVSDYTEVDNNV
TRHLDKVLKRGDWDILILHYLGLDHIGHISGPNSPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTCCAGTTATCGTGACGCACCTTGAAAGTCTGAGAGCTACTGCCCTACA
GAAAGTTACTAGTGCCTAAAGCTGGCGCTGGCACTGATGTTACTGCTGCTGGAGTACA
ACTCCCTATAGAAAACA
CTGCCAGCACCTTAAGACCACTCACACCTCAGAGTGAAGAACTTAAAC
CCGAAGAAATTCA
GCATTCATGACCAGGATCACAAAGTACTGGCCTGGACTCTGGGAATCTCAT
AGCAGTTCCAGATAAAA
ACTACATACGCCAGAGATCTTCTTGCA
TTAGCCTCATCCTTGAGCT
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGTCTCTAAAGGGAGTTTGCTCTAC
TGTGACAAGGATAAAGGACAAAGTCATCCATCCCTCAGCTGAAGAAGGAGAA
ACTGATGAAGCT
GGCTGCCAAAAGGAATCAGCACGCCGCCCTCATCTTATAGGGCTCAGGTGGCCTGGA
ACATGCTGGAGTCGGGGCTCACCCGGATGGTCATCTGCACCTCCTGCAATTGTAATGAGCCT
GTTGGGGTGACAGATAAATTGAGAACAGGAAACACATTGAATTTCATTCAACCAGTTGCAA
AGCTGAAATGAGCCCCAGTGAGGTCA
GGCAGTAGGAACTGCCCTTGAAACGCCCTCGCTA
ATTGAACTAATTGTATAAAA
ACCAACCTGCTCACT

FIGURE 142

MLLLLLEYNFPPIENNQHLKTTHTFRVKNLNPKKFSIHQDHKVLVLDGNLIAVPDKNYIRPEI
FFALASSLSSASAEGSPILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAQKESARRPFI
FYRAQVGWSWNMLESAAHPGWFICTSCNCNEPVGVTDFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 33-36

N-myristoylation site.
amino acids 50-55, 87-92

Interleukin-1
amino acids 37-182

FIGURE 143

CTAGAGAGTATAGGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCTCCCTTAA
TCCAGGATCCTGTCCTCCTGTCCTGTAGGAGTGCCTGTTGCCAGTGTGGGTGAGACAAGTTG
TCCCACAGGGCTGTCTGAGCAGATAAGATTAAGGGCTGGGTCTGTGCTCAATTAACTCCTGTGG
CACGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCAATGCTGGCCTGCCGTGG
AAGGGAGGTCTGTCCTGGCGCTGCTGCTCTTAGGCTCCCAGATCCTGCTGATCTATGC
CTGGCATTCCACGAGCAAAGGGACTGTGATGAACACAATGTCATGGCTCGTTACCTCCCTGCCA
CAGTGGAGTTGCTGTCCACACATTCAACCAACAGAGCAAGGACTACTATGCCACAGACTGGGG
CACATCTTGAATTCTGGAAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT
GGGGAGAACTAGGTGTGGAAATTGAAGACGACATTGACAACGCCATTCCAAGAAAGCACAG
AGCTGAACAATACTTCACCTGCTTCAACCATCAGCACCAAGGCCCTGGATGACTCAGTCAGC
CTCCTGAACAAGACCTGCTGGAGGGATTCCACTTGAGTGAAACCCACTCACAGGTTGTCCATGT
GCTGCTCCCACATTCCGTGGACATCAGCACTACTCTCCTGAGGACTCTCAGTGGCTGAGCAGCT
TTGGACTTGTGTTATCCTATTTGCATGTGTTGAGATCTCAGATCAGTGTGTTAGAAAATCC
ACACATCTTGAGCCTAATCATGTAGTGTAGATCATTAAACATCAGCATTAAAGAAAAAAAAAAA
AA

FIGURE 144

MLGLPWKGGLSWALLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTISTRP
WMTQFSLLNKTCLEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGGCCAAGGTTCTGACGCGATGAGGAAG
CACCTGAGCTGGTGGCTGCCACTGTCTGCATGCTGCTTCAGCCACCTCTCGCGTCCA
GACGAGGGGCATCAAGCACAGAATCAAGTGGAACCGGAAGGCCCTGCCAGCACTGCCAGATCA
CTGAGGCCAGGTGGCTGAGAACCGCCGGAGCCTCATCAAGCAAGGCCAGCTGACATT
GACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTCCCCGATGGCATCCA
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTGTACCGGCTGCATCAATGCCA
CCCAGGCGCGAACCAAGGGGAGTTCCAGAACAGCAGAACAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTGGAGAGGGCGCAGGACT
TCGGGTCAACCATGCACCAGCCAGTGCTCCTCTGCCTTCTGGCTTGATCTGGCTCATGGTAAA
AAGCTTGCCAGGGCTGCAGTACCCAGT
CTTCTCCCCAAACCCACCGTGTCTGAAGGTGCCAGGAGCGCGATGCACTCGCACTGCAA
TGCCGCTCCACGTATGCGCCCTGGTATGTGCCTGCGTTCTGATAGATGGGGACTGTGGTTCT
CCGTCACTCCATTCTCAGCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGTAAATGCTTGAT
GAGAAGAACACATCAGGCACTGCGCCACCTGCTCACAGTACTTCCAACAACACTTAGAGGTAG
GTGTATTCCGTTTACAGATAAGGAAACTGAGGCCAGAGAGCTGAAGTACTGCACCCAGCATT
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCCTGGTTGCTAACCCAGGTTCTGCT
GTCCAATTCCAGAGCTGTCTGGTGATCACTTATGTCTCACAGGGACCCACATCCAAACATGTAT
CTCTAATGAAATTGTGAAAGCTCCATGTTAGAAATAATGAAAACACCTGA

FIGURE 146

MRKHLWWWLATVCMLLFSHL SAVQTRGIKHRIKWNRKALPSTAQITEAQVAENRPGAFIKQGRK
LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNLHQQ
VLWRLVQELCSLKHC EFWLERGAGLRTMHQPVLLCLLALIWL MVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCAAAGGGCTGGGATTATAGCGTGACCAACATGTCGGTCCAGAGTCTCATTT
CCTGATGATTATAGACTCAAAGAAAACTCATGTTCAGAAGCTCTCTCTGGCCTCCTCT
CTGCTTCTTCCCTTTCTTATTAAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
AAATCTTCATTTGCTTGTCACTGGGGTAGGTCACTGAGCTTAGTTTATTTTGAAATT
CAACTTTCAGATTCAAGGGGTACATGTGAAGGTTGTTATGAGTATATTGCATGATGCTGAGG
TTTGGGGT

FIGURE 148

MFRSSLLFWPPLCLLSLFLLILISSIYSESCKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

FIGURE 149

GTCTCCGCGTCACAGGAACCTCAGCACCCACAGGGCGGACAGCGCTCCCTCTACCTGGAGACTTGAC
TCCCGCGGCCAACCCCTGCTTATCCCTGACCGTCGAGTGTAGAGATCCTGCAGCCGCCAGTCC
CGGCCCTCTCCGCCAACACCCACCCCTCCTGGCTCTTCTGTTTACTCCTCCTTTCATTATA
ACAAAAGCTACAGCTCCAGGAGCCCAGCGCCGGCTGTGACCCAAGCCGAGCGTGGAAAGAATGGGTT
CCTCGGGACCGGCCTGGATTCTGGTTAGTGCTCCGATTCAAGCTTCCCAAACCTGGAGGAA
GCCAAGACAAATCTCTACATAATAGAGAATTAAGTGCAGAAAGACCTTGAATGAACAGATTGCTGAA
GCAGAAGAAGACAAGATTAAAAAAACATATCCTCCAGAAAACAAGCCAGGTAGAGCAACTATTCTT
TGTTGATAACTTGAACCTGCTAAAGGCAATAACAGAAAAGGAAAAATTGAGAAAGAAAGACAATCTA
TAAGAAGCTCCCCACTTGATAATAAGTTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAAGT
ATCGATGATTATGACTCTACTAAGAGTGGATTGGATCATAAATTCAAGATGATCCAGATGGTCTTCA
TCAACTAGACGGGACTCCTTAACCGCTGAAGACATTGTCCATAAAATCGCTGCCAGGATTATGAAG
AAAATGACAGAGCCGTGTTGACAAGATTGTTCTAAACTACTTAATCTCGGCCTTATCACAGAAAGC
CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTACAAAATTAATCTCAAAGGAAGCCAACAA
TTATGAGGAGGATCCAATAAGCCCACAAGCTGGACTGAGAATCAGGCTGGAAAAATACCAGAGAAAG
TGACTCCAATGGCAGCAATTCAAGATGGTCTTGCTAAGGGAGAAAACGATGAAACAGTATCTAACACA
TTAACCTTGACAATGGCTTGGAAAGGAGAACTAAAACCTACAGTGAAGACAACCTTGAGGAACCTCCA
ATATTCCCAAATTCTATGCGCTACTGAAAAGTATTGATTCAAGAAAAGAAGCAAAAGAGAAAGAAA
CACTGATTACTATCATGAAAACACTGATTGACTTTGTGAAGATGATGTTGAAATATGGAACAATATCT
CCAGAAGAAGGTGTTCCCTACCTTGAAACTGGATGAAATGATTGCTCTCAGACCAAAACAAGCT
AGAAAAAAATGCTACTGACAATATAAGCAAGCTTCCCAGCACCACAGAGAAGAGTCATGAAGAAA
CAGACAGTACCAAGGAAGAAGCAGCTAAGATGGAAAAGGAATATGGAAGCTTGAAGGATTCCACAAA
GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAACAGAAGCCTATTGGAAGC
CATCAGAAAAAATTGAATGGTGAAGAAACATGACAAAAGGAAATAAGAAGATTATGACCTT
CAAAGATGAGAGACTTCATCAATAACAGCTGATGCTTATGTGGAGAAAGGCATCCTGACAAGGAA
GAAGCCGAGGCCATCAAGCGCATTATAGCAGCCTGAAAAATGGCAAAGATCCAGGAGTCTTCAA
CTGTTGAGAAACATAATATAGCTAAAACACTTCTAATTCTGTGATTAAAATTTTGACCCAAGG
GTTATTAGAAAGTGTGAATTACAGTAGTTAACCTTTACAAGTGGTAAAACATAGCTTCTTCCC
GTAAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAA
AAAAAAAAAAAAA

FIGURE 150

MGFLGTGTWILVLVLPIQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPG
QSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDFKIVSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEDPNKPTSWTENQAGKIPKVTMMAIQDGLAKGENDETVSNTLTNGL
RRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISPEEGV
SYLENLDEMIALQTKNKNATDNISKLFPAPSEKSHEETDSTKEEAAKMEKEYGSLKDSTKDD
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKKGNKEDYDLSKMRDFINKQADAYVEKGILD
EEAEAIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCCGTGGAGCTCAA
GATGGTCCTGAGTGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGCTTATCTGC
ATAATAACCAGCTCTAGCTGGAGGGCTGCATGCAGGGAAAGGTCTAAAGGTGAAGAGATCAGC
GTGGTCCCCATCGGTGGCTGGATGCCAGCCTGTCATCTGGGTGTCCAGGGTGGAAAG
CCAGTGCCTGTCATGTGGGTGGGCAGGAGCCACTCTAACACTAGAGCCAGTGAACATCATGG
AGCTCTATCTGGTCCAAGGAATCCAAGAGCTTCACCTCTACCCGGCGGACATGGGCTACC
TCCAGCTTCGAGTCGGCTGCCAACCCGGCTGGTCTGTCACGGTGCCTGAAGCCGATCAGCC
TGTCAACTCACCCAGCTCCCGAGAATGGTGGCTGGAATGCCCATCACAGACTTCTACTTCC
AGCAGTGTGACTAGGGCAACGTGCCCCCAGAACCTCCCTGGCAGAGCCAGCTGGGTGAGGGT
GAGTGGAGGAGACCCATGGCGACAATCACTCTCTGCTCTCAGGACCCCCACGTCTGACTTAG
TGGGCACCTGACCACTTGTCTCTGGTCCCAGTTGGATAAAATTCTGAGATTGGAGCTCAGT
CCACGGTCTCCCCACTGGATGGTGTACTGCTGTGGAACCTGTAAAAACCATGTGGGTAAA
CTGGGAATAACATGAAAAGATTCTGTGGGGTGGGGAGTGGTGGGAATCATTCTGCT
TAATGGTAACTGACAAGTGTACCCCTGAGCCCCGAGGCCAACCCATCCCAGTTGAGCCTTATA
GGGTCACTAGCTCCACATGAAGTCCTGCACTCACCAGTGCAGGAGAGGGAGGTGGTCATA
GAGTCAGGGATCTATGGCCCTGGCCAGCCCCACCCCTCCCTTAATCTGCCACTGTCT
TGCTACCTTCTATCTCTCCATCATCTTGTGGCATGAGGAGGTGGTGTAGTCAGAA
GAAATGGCTCGAGCTCAGAAGATAAAAGATAAGTAGGGTATGCTGATCCTTTAAAAACCAA
GATACAATCAAATCCCAGATGCTGGTCTCTTCAATTGCTGACATATTGAGA
AGACCTACTTACAAAGTGGCATATATTGCAATTATTAAATTAAAAGATACTATTATATT
TCTTATAGAAAAAGTCTGGAAGAGTTACTCAATTGTAGCAATGTCAGGGTGGCAGTAT
AGGTGATTTCTTTAATTCTGTTAATTATCTGTATTCTCTAATTGAGATAAGATGA
ATTCTTGATAAAAATAAGAAAAGAAATTATCTGAGGTAAGCAGAGCAGACATCATCTGA
TTGTCCTCAGCCTCCACTCCCCAGAGTAAATTCAAATTGAATCGAGCTGCTGCTGGTGG
TTGTAGTAGTGTAGCAGGAAACAGATCTCAGCAAAGCCACTGAGGAGGAGGTGTGAGTTGT
GTGGCTGGAATCTCTGGTAAGGAACCTAAAGAACAAAATCATCTGGTAATTCTTCTAGAAG
GATCACAGCCCCCTGGGATTCCAAGGCATTGGATCCAGTCTCAAGAAGGCTGCTGTACTGGTGA
ATTGTGTCCCCCTCAAATTCACATCCTCTGGAAATCTCAGTCTGTGAGTTATTGGAGATAAG
GTCTCTGCAGATGTAGTTAGTAAGACAAGGTATGCTGGATGAAGGTAGACCTAAATTCAATAT
GACTGGTTCTGTATGAAAAGGAGAGGACACAGAGACAGAGGAGACGCGGGGAAGACTATGTA
AAGATGAAGGCAGAGATGGAGTTGCAGCCACAAGCTAAGAAACACCAAGGATTGGCAACC
ATCAGAAGCTGGAGAGGCAAAGAAGAATTCTCCCTAGAGGCTTAGAGGGATAACGGCTCTG
CTGAAACCTTAATCTCAGACTTCCAGCCTCTGAACGAAGAAAGAATAAATTGGCTGTTAA
GCCACCAAGGATAATTGGTTACAGCAGCTCTAGGAAACTAATACAGCTGCTAAATGATCCCTGT
CTCCTCGTGTTCACATTCTGTGTGCTCCCTCCACAATGTACCAAAGTTGTCTTGTGACCAA
TAGAATATGGCAGAAGTGTGATGGCATGCCACTTCCAAGATTAGGTTATAAAAGACACTGCAGCTC
TACTGAGCCCTCTCTGCCCCACCGCCCCAATCTATCTGGCTACTCGCTCTGGGG
AAGCTAGCTGCCATGCTATGAGCAGGCCTATAAAGAGACTTACGTGGTAAAAAAATGAAGTCTCT
GCCACAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATCGATGTTGTGTT
AAGTGCTCAGTTGGTCTAACTGTTATGCAGCAATAGATAAATAATGCAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEIISVVPNRWLDASLSPVILGVQGGS
QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP
VRLTQLPENGGWNAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCAATGGCCGC
CCTGCAGAAATCTGTGAGCTTTCTTATGGGGACCCTGCCACCAGCTGCCCTCTCTGG
CCCTCTTGGTACAGGGAGGAGCAGCTGCCCATCAGCTCCACTGCAGGCTTGACAAGTCCAAC
TTCCAGCAGCCCTATATCACCAACCGCACCTCATGCTGGCTAAGGAGGCTAGCTGGCTGATAA
CAACACAGACGTTCGTCTCATTGGGGAGAAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACCTCACCCCTGAAGAAGTGCTGTTCCCTCAATCTGATAGGTT
CAGCCTTATATGCAGGAGGTGGTGCCTGCCCTGCCAGGCTCAGCAACAGGCTAAGCACATGTCA
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGAAAAGCTGAAGGACACAGTGAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTTGCTGTTATGTCTGAGAAAT
GCCTGCATTTGACCAGAGCAAAGCTGAAAATGAATAACTAACCCCCTTCCCTGCTAGAAATAA
CAATTAGATGCCCAAAGCGATTTTAACCAAAAGGAAGATGGGAAGCCAACTCCATCATG
ATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGAAACCAATGCCACTTTGTTATA
AGACCAGAAGGTAGACTTCTAAGCATAGATATTATTGATAACATTGTAACTGGTGTTC
TATACACAGAAAACAATTATTTAAATAATTGTCTTTCCATAATCAGTACTTATATTATAAA
TCCTTTAGGGAAAAAACCCCTAAATAGCTCATGTTCCATAATCAGTACTTATATTATAAA
TGTATTATTATTATAAGACTGCATTTATTATCATTATTAAATATGGATTTATTAT
AGAACATCATTGATATTGCTACTTGAGTGTAAAGGCTAATATTGATATTATGACAATAATTAT
AGAGCTATAACATGTTATTGACCTCAATAAACACTTGGATATCCC

FIGURE 154

MAALQKSVSSFLMGTLATSCLLLALLVQGGAAAPISHCRLDKSNFQQPYITNRTFMLAKEASL
ADNNNTDVRLIGEKLFHGVMSERCYLMKQVLNFTLEEVLFQSDRFQPYMQEVVPFLARLSNRLS
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTCCACGAGGCCTGT
CAGTCAGTGCCCGACTTGTGACTGAGTGTGCAGTGCCAGCATGTACCAGGTAGTCAGAGGGC
TGCCTGAGGGCTGTGCTGAGAGGGAGAGGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCCAAGC
TGCCAGGTTGGGCTGGGGCCAAGTGGAGTGAGAAACTGGATCCAGGGGAGGGTGCAGAT
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCATAGCCTTCTACAGGTGGTTGCAT
TCTTGGCAATGGTCATGGGAACCCACACCTACAGCCACTGGCCAGCTGCTGCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCCTCCCTAGAGCCTGCTAG
GCCCAACCGCCACCCAGAGTCCTGTAGGCCAGTGAAGATGGACCCCTAACAGCAGGGCCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCCAGGACCTGTACACGCCGT
TGCCTGTGCCGCACTGCGTCAGCCTACAGACAGGCTCCACATGGACCCCCGGGCAACTCGGA
GCTGCTCTACCACAAACCAGACTGTCTTCTACAGGCCATGCCATGGCGAGAAGGGCACCCACA
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTAGCTTGTGTGCGGGCCCGT
GTGATGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTTTGGAAACCTGGAGCCAGGTGTACA
ACCACTGCCATGAAGGGCAGGATGCCAGATGCTTGGCCCTGTGAAGTGTGTGCGAGCAG
CAGGATCCCAGGACAGGATGGGGCTTGGGAAAACCTGCACTCTGCACATTGAAAAGAG
CAGCTGCTGCTTAGGCCGCCGAAGCTGGTGTCTGTCACTTCTCAGGAAAGGTTCAA
GTTCTGCCATTCTGGAGGCCACCACTCCTGTCTTCTCTTCCATCCCTGCTACCTG
GCCAGCACAGGCACCTTAGATATTCCCCCTGCTGGAGAAGAAAGGCCCTGGTTTATT
TGTTGTTACTCATCACTCAGTGAGCATCTACTTGGTGCATTCTAGTGTAGTTACTAGTCTT
TTGACATGGATGATTCTGAGGAGGAAGCTGTTATTGAATGTATAGAGATTATCAAATAATAT
CTTATTTAAAAATGAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGHTYSHWPSCCPKGQDTSEELLRWSTPVPPLEPA
RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTHKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCG**ATGT**CGCTCGTGTGCTAACGCCCTGGCCGCGTGTGCAGGAGCGCCGTACCCCGAGAGCC
GACCGTTCAATGTGGCTCTGAAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCC
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACACTAGTGGTGAACAGGGGACTATTCA
ATTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCCGCTTGTGAAGGCCACCAA
GATTGTGTGACGGGAAAAGCAACTTCCAGTCCTACAGCTGTGAGGTGCAATTACACAGAGG
CCTTCCAGACTCAGACCAGACCCCTGGTGTAAATGGACATTTCTACATCGGCTCCCTGTA
GAGCTGAACACAGTCTATTCATTGGGCCATAATATTCTAATGCAAATATGAATGAAGATGG
CCCTTCCATGTCTGTGAATTCACCTCACCAAGGCTGCCTAGACCACATAATGAAATATAAAAAAA
AGTGTGTCAAGGCCGAAGCCTGTGGGATCCGAACATCACTGTTGTAAGAAGAATGAGGAGACA
GTAGAAAGTGAACCTCACAACCCTCCCCTGGAAACAGATACTGGCTTTATCCAACACAGCAC
TATCATGGGTTTCAGGTGTTGAGCCACACCAGAAGAAACAAACGCGAGCTCAGTGGTGA
TTCCAGTGACTGGGATAGTGAAGGTGCTACGGTGAGCTGACTCCATATTCCTACTGTGGC
AGCGACTGCATCCGACATAAGGAACAGTTGTCCTGCCACAAACAGGCGTCCCTTCCCTCT
GGATAACAACAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTCTGCTCTGCTGGTGGCCA
CATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTTCTT
TCTACCACCACTACTGCCCTTCAAGGTTCTGTGGTTACCCATCTGAAATATGTTCCA
TCACACAATTGTTACTTCACTGAATTCTCAAAACCATTGAGGTGAGGTGATCCTTGAAA
AGTGGCAGAAAAGAAAATAGCAGAGATGGGTCCAGTGCAGTGGCTGCCACTAAAAGAAGGCA
GCAGACAAAGTCGTCTCCTTCCATGACGTCAACAGTGTGCGATGGTACCTGTGGCAA
GAGCGAGGGCAGTCCAGTGAGAACTCTCAAGACCTCTCCCCCTGCCTTAACCTTCTGCA
GTGATCTAAGAAGCCAGATTCTGACAAATACGTGGTGGTCTACTTAGAGAGATTGATACA
AAAGACGATTACAATGCTCTCAGTGTCTGCCCAAGTACCACTCATGAAGGATGCCACTGCTT
CTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCCACG
ATGGCTGCTGCTCTTGT**AG**

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLIPGDLRDLRVEPVTTSVATGDYSILMNVS WV
LRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVELNTVYFIGAHNIP
NANMNEDGPSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKNEETVEVNFTTPLGNRYMALIQH
STIIGFSQVFEPHQKKQTRASVVIPTVGDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPQTGVPFPLDNNK
SKPGGWLPLLLSSLVATWVLVAGIYLMWRHERIKKTSFSTTLLPPIKVLVVYPSEICFHHTICYFTEFL
QNHCRCSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVFLSNDVNSVCDGTCGKSEGSPSENSQDLFPLA
FNLFCSDLRSQIHLHKVVVYFREIDTKDDYNALSVCPKYHLMKDATAFCAELLHVKQQVSAGKRSQACHD
GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283 - 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

FIGURE 159

AGCCACCAGCGAACATGACAGTGAAGACCCTGCATGGCCCAGCCATGGTCAAGTACTTGCTGCT
GTCGATATTGGGGCTTGCCTTCTGAGTGAGGCAGCTCGGAAAATCCCCAAAGTAGGACATA
CTTTTTCCAAAAGCCTGAGAGTTGCCCGCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
ATCATCAATGAAAACCAGCGCGTTCCATGTCACGTAACATCGAGAGCCGCTCCACCTCCCCCTG
GAATTACACTGTCACTTGGGACCCCAACCGTACCCCTCGGAAGTTGTACAGGCCAGTGTAGGA
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTTCCATCCAGCAA
GAGACCCTGGTCGTCCGGAGGAAGCACCAAGGCTGCTCTGTTCTTCCAGTTGGAGAAGGTGCT
GGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCATCCACCATGTGCAGTAAGAGGTGCATATCC
ACTCAGCTGAAGAAG

FIGURE 160

MTVKTLHGPAMVKYLLLSILGLAFLSEAAARKIPKGHTFFQKPESCPPVPGGSMKLDIGIINEN
QRVMSMRNIESRSTSPWNYTVTWDPNRYPSEVVQAQCRNLGCINAQGKEDISMNSVPIQQETLVV
RRKHQGCSVSFQLEKVLVTVGCTCVTPVIHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

ACACTGGCCAAACAAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGAGTCAGGACTCCCCAGG
ACAGAGAGTGCACAAACTACCCAGCACAGCCCCCTCCGCCCCCTGGAGGCTGAAGAGGGATT
CAGCCCCCTGCCACCCACAGACACGGGCTGACTGGGGTGTCTGCCCTTGGGGGGGGCAGCAC
AGGGCCTCAGGCTGGGTGCCACCTGGCACCTAGAAGATGCCTGTGCCCTGGTCTTGCTGTCCT
TGGCACTGGGCCGAAGCCCAGTGGCCTTCTCTGGAGAGGCTTGTGGGCCCTCAGGACGCTACC
CACTGCTCTCCGGGCTCTCCTGCCCTCTGGGACAGTGACATACTCTGCCCTGCCCTGGGACAT
CGTGCCTGCTCCGGGCCCCGTGCTGGGCCCTACGCACCTGCAGACAGAGCTGGTGTGAGGTGCC
AGAAGGAGACCGACTGTGACCTCTGTCTGCGTGTGGCTGTCCACTTGGCGTGCATGGCACTGG
GAAGAGCCTGAAGATGAGGAAAAGTTGGAGGAGCAGCTGACTCAGGGTGGAGGAGCCTAGGAA
TGCCTCTCCAGGCCAAGTCGTGCTCTCCTCCAGGCTACCCACTTGGCCCTGCCCTGCGTGC
TGGAGGTGCAAGTGCCTGCTGCCCTTGTGCAGTTGGTCAGTCTGTGGCTCTGTGGTATATGAC
TGCTTCGAGGCTGCCCTAGGGAGTGAGGTACGAATCTGGCCTATACTCAGGCCAGGTACGAGAA
GGAACACTAACACACACAGCAGCTGCCCTGCCCTGGCTAACGTGTACGCAGATGGTAC
ACGTGCATCTGGTTCTGAATGTCTTGAGGAGCAGCACTCGGCCCTCCCTGTACTGGAATCAG
GTCCAGGGCCCCCAAAACCCCGGTGGCACAAAACCTGACTGGACCGCAGATCATTACCTTGAA
CCACACAGACCTGGTCCCTGCCCTGTATTAGGTGTGGCCTCTGGAACCTGACTCCGTTAGGA
CGAACATCTGCCCTTCAGGGAGGACCCCGCGCACACCAAGAACCTCTGGCAAGGCCCGACTG
CGACTGCTGACCCCTGCAGAGCTGGCTGGACGCACCGTGCTGCCCTGGCAGAACGGCACT
GTGCTGGCGGGCTCCGGGTGGGGACCCCTGCCAGCCACTGGTCCACCGCTTCCCTGGGAGAACG
TCACTGTGGACAAGGTTCTCGAGTCCATTGCTGAAAGGCCACCCCTAACCTCTGTGTTAGGTG
AACAGCTGGAGAAGCTGCAGCTGCAGGAGTGCAGCTGTGGCTGACTCCCTGGGCCTCTCAAAGA
CGATGTGCTACTGTTGGAGACACGAGGCCCCCAGGACAACAGATCCCTCTGTGCCTTGGAACCC
GTGGCTGTACTTCACTACCCAGCAAAGCCTCACGAGGGCAGCTGCCCTGGAGAGTACTTACTA
CAAGACCTGCAGTCAGGCCAGTGTCTGCAGCTATGGGACGATGACTTGGAGCGCTATGGCCTG
CCCCATGGACAAATACATCCACAAGCGCTGGCCCTCGTGTGGCTGGCCTGCCTACTTTGCCG
CTGCCTTCCCTATCCTCTCAAAAGGATCACCGAAGGGTGGCTGAGGCTTGTGAAA
CAGGACGTCCGCTCGGGGGCGCCAGGGCGCGCTGTGCTCCCTACTCAGCCGATGA
CTCGGTTTCGAGCGCCTGGTGGCGCCCTGGCGTGGCCCTGTGCCAGCTGCCGTGCGCTGG
CCGTAGACCTGTGGAGGCCGTGAACCTGAGCGCGCAGGGGCCGTGGCTTGGTTACGCGCAG
CGGCCAGACCCCTGCAGGAGGGCGCGTGGTGGCTTGTCTCTCCCGTGCCTGGCGCT
GTGCAGCGAGTGGCTACAGGATGGGTGTCCGGCCGGCGCACGGCCCGACGACGCCCTCC
GCCCTCGCTCAGCTCGTGTGCCGACTTCTGCAGGGCCGGCGCCCGCAGCTACGTGGGG
GCCCTCGACAGGCTGCTCACCCGGACGCCGTACCCGCCCTTCCGCACCGTGCCCGTCTT
CACACTGCCCTCCCAACTGCCAGACTTCTGGGGCCCTGCAGCAGCCTCGGCCCGTCCG
GGCGGCTCCAAGAGAGAGCGGAGCAAGTGTCCCGGGCCCTCAGCAGCCCTGGATAGCTACTC
CATCCCCGGGACTCCCGGCCGGACGCCGGGTGGGACCAGGGCGGGACCTGGGGGGGG
CGGGACTTAAATAAAGGCAGACGCTGTTTCTAAAAAA

FIGURE 162

MPVPWFLLSLALGRSPVVLSLERLVPQDATHCSPGLSCLRWDSDILCLPGDIVPAPGPVLAPTHLQTELV
LRCQKETDCDLCLRVAVHLAVHGHWEEPEDEEKFGGAADSGVEEPRNASLQAQVVL SFQAYPTARCVLLEV
QVPAALVQFGQSVGSVYDCFEAALGSEVRIWSYTQPRYEKELNHTQQLPALPWLNVSADGDNVHLVLNVS
EEQHFGLSLYWNQVQGPPKPRWHKNLTGPQIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQN
LWQAARLRLTLQSWLLDAPCSLPAEAALCWRAPGGDPCQPLVPPLSWENVTVDKLEFPLLKGHPNLCVQ
VNSSEKLQLQECILWADSLGPLKDDVLLLETRGQDNRSILCALEPSGCTSLPSKASTRAARLGEYLLQDIQS
GQCLQLWDDDLGALWACPMDKYIHKRWALVWLACLLFAAALSLLLLKKDHAKGWLRLKQDVRSGAAARG
RAALLLYSADDSGFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAWFHAQRRTLQEGGVVLLFSP
GAVALCSEWLQDGVS GPGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLLHPDAVPALFRTVPVFT
LPSQLPDFLGALQQPRAPRSGRLQERAEQVSRALQPALDSYFHPPGTPAPGRGVGPAGPGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,
334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,
692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATCTGACTGTGGGATCCCTGGCT
GCTCACGCCCTGAGGACCCCTCGGATCTGCTCAGCACGTGAAATTCCAGTCCAGCAACTTGA
AAACATCCTGACGTGGACAGCAGGGCCAGAGGGCACCCAGACACGGTCTACAGCATCGAGTATA
AGACGTACGGAGAGAGGGACTGGGTGGCAAAGAAGGGCTGTCAAGGGATCACCCGGAAAGTCCTGC
AACCTGACGGTGGAGACGGCAACCTCACGGAGCTCTACTATGCCAGGGTACCGCT
GTCAGTGCAGGGAGGCCGGTCAGCCACCAAGATGACTGACAGGTTCAAGCTCTGCAGCACACTAC
CCTCAAGCCACCTGATGTGACCTGTATCTCAAAGTGAGATCGATTCAAGATGATTGTTCATCCTA
CCCCCACGCCAATCCGTGCAGGCCATGGCTAACCCCTGGAAGACATCTCCATGACCTG
TTCTACCACTTAGAGCTCCAGGTCAACCGCACCTACCAATGCACCTGGAGGGAAAGCAGAGAGA
ATATGAGTTCTCGCCCTGACCCCTGACACAGAGTTCTGGCACCATCATGATTGCGTTCCA
CCTGGGCCAAGGAGAGTGCCTTACATGTGCCAGTGAAGACACTGCCAGACGGACATGGACC
TACTCCTTCTCCGGAGCCTTCTGTTCTCCATGGGCTTCAGTGCAGTACTCTGCTACCTGAG
CTACAGATATGTCACCAAGGCCCTGCACCTCCAACTCCCTGAACGTCCAGCGAGTCCTGACTT
TCCAGCCGCTGCCTCATCCAGGAGCACGTCTGATCCCTGTCTTGACCTCAGCGGCCAGC
AGTCTGGCCCAGCCTGTCCAGTACTCCAGATCAGGGTGTCTGGACCCAGGGAGCCCGCAGGAGC
TCCACAGCGGCATAGCCTGTCCAGTACCTACTTAGGGCAGCCAGACATCTCCATCCTCCAGC
CCTCCAACGTGCCACCTCCCCAGATCCCTCCCCACTGTCTATGCCCAAACGCTGCCCTGAG
GTCGGGCCCATCCTATGCACCTCAGGTGACCCCCGAAGCTCAATTCCATTCTACGCCCA
GGCCATCTCTAAGGTCAGCCTCCATGCCCTCAAGCCACTCCGGACAGCTGGCCTCC
CCTATGGGTATGCATGGAAGGTTCTGGCAAAGACTCCCCACTGGGACACTTCTAGTCCTAAA
CACCTTAGGCCTAAAGGTCAAGCTCAGAAAGAGCCACCAGCTGGAGCTGCATGTTAGGTGGCCT
TTCTCTGCAGGAGGTGACCTCTGGCTATGGAGGAATCCAAGAAGCAAATCATTGCACCAGC
CCCTGGGATTGACAGACAGAACATCTGACCCAAATGTGCTACACAGTGGGAGGAAGGGACA
CCACAGTACCTAAAGGCCAGCTCCCCCTCCTCAGTCCAGATCGAGGGCCACCCATGTC
CCTCCCTTGCAACCTCCTCCGGTCCATGTTCCCCCTCGGACCAAGGTCCAAGTCCCTGGGCC
TGCTGGAGTCCCTTGTGTCTGAGGATGAAGCCAAGAGGCCAGCCCCCTGAGACCTCAGACCTG
GAGCAGCCCACAGAACTGGATTCTCTTCAGAGGCCTGGCCCTGACTGTGCAGTGGAGTCCTG
AGGGGAATGGGAAAGGCTTGGTGCTTCCTCCCTGTCCCTACCCAGTGTACATCCTGGCTGTCA
ATCCCATGCCCTGCCATGCCACACACTCTGCGATCTGGCCTCAGACGGGTGCCCTTGAGAGAAC
AGAGGGAGTGGCATGCAGGGCCCTGCCATGGGTGCGCTCCTCACCGGAACAAAGCAGCATGATA
AGGACTGCAGGGGGAGCTCTGGGAGCAGCTGTGAGACAAGCGCGTGCCTGAGCCCTG
CAAGGCAGAAATGACAGTCAAGGAGGAATGCAGGGAAACTCCCGAGGTCCAGAGGCCACCTC
CTAACACCATGGATTCAAAGTCTCAGGGAAATTGCCTCTCCTGCCCTTCCTGGCCAGTT
ACAATCTAGCTCGACAGAGCATGAGGCCCTGCCTCTCTGCTATTGTCATGAAAGGAGGAGA
GCCTGGAAAAGAACCAAGGCCCTGGAAAAGAACCAAGAAGGAGGCTGGCAGAACAGAACAC
ACTTCTGCCAAGGCCAGGGCCAGCAGGACGGCAGGACTCTAGGGAGGGTGTGGCCTGCAGCTCA
TTCCCAGCCAGGGCAACTGCCTGACGTTGACGATTTCAGCTTCATTCTGATAGAACAAAGC
GAAATGCAGGTCCACCAGGGAGGGAGACACACAAGCCTTTCTGCAGGCAGGAGTTCA
ATCCTGAGAATGGGTTGAAAGGAAGGTGAGGGCTGTGGCCCTGGACGGGTACAATAACAC
TGTACTGATGTCACAACATTGCAAGCTCTGCCTGGGTTCAAGCCATCTGGCTCAAATTCCAGC
CTCACCACTACAAGCTGTGACTCAAACAAATGAAATCAGTGCCTGAGAACCTCGGTTCTC
ATCTGTAATGTGGGATCATAACACCTACCTCATGGAGTTGTGGTGAAGATGAAATGAAGTCATG
TCTTAAAGTGTAAATAGTGCCTGGTACATGGCAGTGCCAATAACGGTAGCTATTAAAAA
AAAAAAA

FIGURE 164

MRTLLTILTGVSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGPTDTVYSIEYKTYGERDW
VAKKGCQRITRKSCNLTVETGNLTELYYARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDVT CIS
KVRSIQMIHVPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHLGGKOREYEFFGLTPDT
EFLGTIMICVPTWAKESAPYMCRVKTLPDRTWTYSFSGAFLFSMGFLVAVLCYLSRYVTKPPAP
PNSLNVQRVLTFQPLRFIQEHVLIPIVFDLSSGPSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT
YLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFPFYAPQAISKVQPSSY
APQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPKGQLQKEPPAGSCMLGGSLQEVTSIAM
EESQEAKSLHQPLGICTDRTSDPNVLHSGEEGTPQYLGQLPLLSSVQIEGHPMSLPLQPPSGPC
SPSDQGPSPWGLLESILVCPKDEAKSPAETSDLEQPTELDSLFRGLALTQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAGTCACCCGGGCCGGTGGCCACAACATGG
CTGCGGCGCCGGGCTGCTCTCTGGCTGTCGTGGGGCGCTCTGGTGGTCCCAGGCCAG
TCGGATCTCAGCCACGGACGGCGTTCTGGACCTCAAAGTGTGCGGGGACGAAGAGTCAGCAT
GTTAATGTACCGTGGAAAGCTCTGAAGACTTCACGGGCCCTGATTGTCGTTGTGAATTAA
AAAAAGGTGACGATGTATGTCTACTACAAACTGGCAGGGGATCCCTGAACCTTGGCTGGA
AGTGTGAACACAGTTGGATATTCCAAAAGATTGATCAAGGTACTTCATAAAACACGGA
AGAAGAGCTACATATTCCAGCAGATGAGACAGACTTGTCTGCTTGAAGGAGGAAGAGATGATT
TTAATAGTTATAATGTAGAAGAGCTTTAGGATCTTGAACTGGAGGACTCTGTACCTGAAGAG
TCGAAGAAAGCTGAAGAAGTTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCGGGGCGTGA
ACTTGACCTGTGCCTGAGCCGAGGCATTAGCTGAGGACTCTGAGGAGATGGAGAAGGTGCTTCT
CAGAGAGCACCAGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGCCACCCACACCCAGCGGT
CCTGCGGCTAACGCTCAGGGAGTCAGTCTCGTGGACACTTGAAGAAATTCTGCACGATAA
ATTGAAAGTGCCGGAAAGCAGAACTGGCAATAGTTCTCCTGCCTCGTGGAGCAGTCGTTATT
AGACAGATGCTTACAAAGTCCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTCGTTATT
CATTACAGCAAAGGATTCGTTGGCATAAAATCTAAGTTGTTACAAAGATTGTTTTAGTA
CTAAGCTGCCTGGCAGTTGCATTTGAGCAAACAAAATATTATTCCCTTAAGTA
AAAAAA

FIGURE 166

MAAAPGLLFWLFVLGALWWVPGQSDLSHGRRFSDLKVC~~GDEEC~~SMLMYRGKALEDFTGPDCRFVN
FKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEEELH~~I~~PADETDFVC~~FEGGRD~~
DFNSYNVEELLGSLELED~~S~~VPEESKKAAEVSQHREKSPEESRGRELD~~P~~VPEPEAFRADSE~~D~~GE~~G~~
FSESTEGLQGQPSAQESH~~P~~H~~T~~SGPAANAQGVQSSLDTFEEILHDKLK~~V~~PGSESRTGNSSPASVER
EKT~~D~~AYKVLKTEMSQRGSGQCVIHYSKGFRWHQNL~~S~~FYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGGACCAGGGCGCACCGGCTCAGCCCTCTCACTTGTCAAGAGGCCGGGAAAGAGAAGCAAAGCGC
AACGGTGTGGTCCAAGCCGGGCTTCTGCTTCGCTCTAGGACATACACGGGACCCCTAACTTC
AGTCCCCAAACGCGCACCTCGAAGTCTGAACCTCCAGCCCCCACATCCACGCGCGGACAGG
CGCGGCAGGCGGCAGGTCCCAGGCGAAGGCGATGCGCGCAGGGGTGGCAGCTGGCTCGGG
GGCAGGGAGTAGGGCCGGCAGGGAGGCAGGGAGGCTGCATATTCAAGAGTCGCGGGCTGCGCC
GGCAGAGGCCGCGCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCGCAGTGA
TCTCGCTGCTGGCGCCGCGCTGCTCTGCCACGGAGCCTCTGCCGCCGCGTGGTCAGC
GGCCAAAAGGTGTGTTGCTGACTCAAGCATCCCTGCTACAAAATGGCTACTTCCATGA
GTCCAGCCGAGTGAAGCTTCAAGGAGGCACGCCCTGGCTGAGAGTGAGGGAGGAGTC
GCCTTGAGAATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGCAAAACCTGACA
ACAGGGATTCTGATGGTATTCTGGATAGGGCTTGGAGGAATGGAGATGGCAAACATCTGG
TGCCTGCCAGATCTTACCAAGTGGTCTGATGGAAGCAATTCCAGTACCGAAACTGGTACACAG
ATGAACCTTCTGCGGAAGTGAAGAAGTGTGTTGATGACCAACCAACTGCCAATCCTGGC
CTTGGGGTCCCTACCTTACCAAGTGGAATGATGACAGGTGTAACATGAAGCACAATTATTTG
CAAGTATGAACCAGAGATTAATCCAACAGCCCTGTAGAAAAGCCTTATCTACAAATCA
GAGACACCCATCAGAATGTGGTGTACTGAAGCAGGTATAATTCCAATCTAATTATGTT
ATACCAACAATACCCCTGCTTACTGATACTGGTGCTTGGAACCTGTTCCAGATGCT
GCATAAAAGTAAAGGAAGAACAAAAACTAGTCAAACACCAGTCTACACTGTGGATTCA
CCAGAAAAGAAAGTGGCATGGAAGTATAAACTCATTGACTGGTCCAGAATTGTAATTCT
GGATCTGTATAAGGAATGGCATCAGAACAAATAGCTGGAATGGCTGAAATCACAAAGGATCTGC
AAGATGAACTGTAAGCTCCCCCTGAGGCAAATATTAAAGTAATTGTTATATGCTATT
TTAAAGAATATGCTGTGCTAATAATGGAGTGAGACATGCTATTGCTAAAGGATGC
ACTTCAAACCTCAAGCAAATGAAATGGACAATGCGAGATAAAGTTGTTATCAACAC
GTGCTTGGCTTTTATTCTTACCTTCAAGTGTGCTTACACTGTGAAAGTATT
TGTATATTGATTGAAATTACAGTGTGCAAAGTATTACCTTGCATAAGTGTGATA
ATGAACACTGTCTAATATTATTATGGCATCTCATTTCATAACATGCTCTTGATTAAAG
AAACTTATTACTGTTGCAACTGAATTCACACACACACAAATATAGTACCATAG
TTTCTCGAAATAATTCACTTCAAGTGTGCTTACACTGTGCTTGGTCAATGTCT
AATAAGAAGCTATTCAATTAGTGTGATATAACCTCCTCAAACATTACTAGAGG
TGTCTAATTCAATTGCAAGACATGTGCTTATAATTATTAGCTTAAACAGATT
TTGTAATAATGTAACTTGTTAATAGGTGCATAAACACTAATGCAGTCATT
TGACATACACAATATAATCATATGTCTCACACGTTGCCTATAATGAGAAG
GGTTCTGAAATCAATGTGGCCCTCTTGCCACTAAACAAAGATGGTGT
ATTGACACTGGAGGCAGATAGTTGCAAAGTTAGTCTAAGGTTCC
ACTATATTAGTATAACAGAGGTATGTGGTTGAGACCAGGTGAATAGT
ACAAGCACAGCACACAGACATTAGGAAGGAAACTACGA
AACCCATCAGTGATCGCATATTCAATTGATGAGGGTTGCTGAG
CTGCTTATCTCTAGTTCTCAATGCTTACGCC
CTGGTCTTCATATGTCCTGTGCTC
AAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAA

FIGURE 168

MSRVVSLLLGAALLCGHGAFCRVVSGQKVCADFKHPCYKMAFHELSSRVSFQEARLACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQTSGACPDLYQWSDGSNSQ
YRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSKGRTKTSPNQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217